The Codex of a Companion of the Prophet and the Qur'ān of the Prophet
Author(s): Behnam Sadeghi and Uwe Bergmann
Source: *Arabica*, T. 57, Fasc. 4 (2010), pp. 343-436
Published by: Brill
Stable URL: https://www.jstor.org/stable/25782619
Accessed: 16-03-2025 02:43 UTC

BRILL

Arabica 57 (2010) 343-436

Arabica

brill.nl/arab

The Codex of a Companion of the Prophet and the Qur'ān of the Prophet

Behnam Sadeghi and Uwe Bergmann[1]

Abstract

The essay discusses a manuscript of the Qur'ān dating from the first half of the seventh century AD. The text does not belong to the 'Uṯmānic textual tradition, making this the only known manuscript of a non-'Uṯmānic text type. The essay compares this text type with those of the 'Uṯmānic and other Companion textual traditions in order to shed light on the Prophetic prototype.

Keywords

Qur'ān, Prophet Muḥammad, 'Uṯmān, Ibn Mas'ūd, Islamic origins, palimpsest

[1] We would like to thank the owner of the Stanford '07 folio for making it available and for submitting a sample for radiocarbon dating, Bryce Cronkite-Ratcliff for help with image tracing, Ceci Evangelista of the Office of Development at Stanford University for logistical and other help, and the staff at Stanford University Libraries and the SSRL. Portions of this research were carried out at the Stanford Synchrotron Radiation Laboratory (SSRL) at the Stanford Linear Accelerator Center (SLAC), a national user facility operated by Stanford University on behalf of the U.S. Department of Energy, Office of Basic Energy Sciences.

Behnam Sadeghi would like to thank the following persons for their written comments on this essay: Michael Cook, David Powers, Patricia Crone, Michael Cooperson, Mohsen Goudarzi, Devin Stewart, Zaid Adhami, and Muhammad Mekki. I also thank Michael Cook and Robert Waltz for helpful conversations and exchanges, Mette Korsholm for providing very high-resolution images of the folio in the David Collection, M.S.M. Saifullah for bibliographical guidance, Scott Lucas for lending his UNESCO CD-ROM, and M.M. al-A'ẓamī for gifting the second edition of his *History of the Qur'ānic Text*. The first version of this essay was presented at the Colloquium on the Early History of the Qur'ān, Stanford University, July 30-31, 2009.

Uwe Bergmann's contributions consisted of creating the instrumentation for and implementing the X-ray fluorescence imaging of the folio called here Stanford '07, determining the composition of the inks on the folio, and participating in preparations for radiocarbon dating. Behnam Sadeghi converted the numerical output of imaging into images of the folio, traced the lower text, calculated one-sided probabilities for the radiocarbon dating results, and wrote this essay. The pronoun "I" in this essay refers to Behnam Sadeghi.

 DOI: 10.1163/157005810X504518

I. Introduction

Preview of Results

This is a study of two Qur'āns: the upper and the lower layers of a palimpsest called here Ṣan'ā' 1. The upper layer of writing, a standard Qur'ān, could be from the first or second half of the seventh century AD, and possibly even early eighth century. Radiocarbon dating assigns the parchment, and hence the lower writing, to the first half of the seventh century.

Early Muslim reports assert that different Companions of the Prophet had different versions of the Qur'ān, and some reports give the purported variants of their codices. The differences among these codices appear to have motivated an attempt at standardization. According to the collective memory of early Muslims, the Companion 'Uṯmān, after becoming caliph, disseminated a version of the holy book, declaring it the standard. The date of this event is uncertain, but it appears to have taken place sometime during AH 24-30, *i.e.* AD 644-650.[2] It is to the textual tradition identified with this version that almost all extant Qur'ānic manuscripts belong.

The main significance of the Ṣan'ā' 1 manuscript is that its lower text does not belong to this 'Uṯmānic textual tradition. In this sense, it is "non-'Uṯmānic." It belongs to some other textual tradition which is designated here as C-1. The C-1 textual tradition is distinct not only from that of 'Uṯmān, which is known from both literary sources and manuscripts, but also from those of Companions Ibn Mas'ūd and Ubayy b. Ka'b, whose recensions of the Qur'ān are not attested in manuscripts, being known only from descriptions in literary sources. I will argue that C-1 and these others formed parallel textual traditions. Comparing them can thus illuminate the state of the text prior to the branching off of these various traditions. It can shed light on the progenitor of all textual traditions, the Qur'ānic prototype.

The essay demonstrates that memory and orality played a role in the genesis of the Companion codices. This fact fully accounts for the differences between the textual traditions of C-1 and 'Uṯmān. The magnitude and number of the differences point to orality. Purely oral transmission, however, is unlikely, since differences among the codices remain the exception rather than the rule, even when it comes to minor elements of language. Nor does a scenario of deliberate redaction of a written text explain the differences. The

[2] For an analysis of the evidence bearing on the date of 'Uṯmān's standardization, see Maḥmūd Rāmyār, *Tārīḫ-i Qur'ān*, 2nd ed., Tehran, Amīr Kabīr, HS 1362/1983, p. 433-5; cf. Ahmad 'Ali al-Imam, *Variant Readings of the Qur'an*, Herndon, Virginia, International Institute of Islamic Thought, 2006, p. 20-1.

evidence calls for a model involving both orality and writing, *i.e.* "semi-orality". In particular, a scenario in which the prototype was dictated provides the appropriate model. The largest differences between the Qur'ānic textual traditions of C-1, 'Uṯmān, etc. go back to when the prototype was recited and taken down by different scribes in somewhat different ways.

This is not to say that the lower writing of Ṣan'ā' 1 itself was the proximate product of dictation. It may have been copied off a manuscript. The point is that the lower writing and its parent manuscript were both part of a "C-1 textual tradition". The C-1 textual tradition and the traditions corresponding to other Companion codices must have branched off at some point. It is this branching off that involved semi-orality. Once the various textual traditions (the standard one and the Companion codices) were born through the dictation of some or all of them, they could have thereafter been handed down through written transmission.

Textual criticism aims to determine a prototype on the basis of different versions of a text. For the purposes of the present task, these versions include the 'Uṯmānic Qur'ān and different Companion codices, and it is thus necessary to discuss their status before undertaking the analysis. I begin with the codex of 'Uṯmān. Drawing out the ramifications of important recent findings (by Michael Cook, Yasin Dutton, Hossein Modarressi, etc.),[3] I synthesize a framework for conceptualizing the relationship of the 'Uṯmānic tradition with non-'Uṯmānic codices and for approaching 'Uṯmānic manuscripts such as the upper text of the palimpsest. As for the status of non-'Uṯmānic codices described in Muslim literary sources, I argue that there must be a significant kernel of truth in these reports, and that a particular second-century list of the variants of the Companion Ibn Mas'ūd may be a largely reliable representation of an early codex.

The first method of textual criticism used is stemmatics, sometimes called the "genealogical method", a procedure based on how frequently each text agrees with each of the others. This approach is applied to three text types—those of 'Uṯmān, C-1, and Ibn Mas'ūd—to construct possible stemmata. The most plausible stemma is shown to be one in which a prototype is the common ancestor of these three Companion codices. Subsequently, the data suggests that among them, the textual tradition of 'Uṯmān gives the most accurate reproduction of the prototype. An alternative scenario explains the data equally well: one may envision 'Uṯmān's codex as a composite formed by

[3] The citations to the works of these scholars and others will follow in this essay in due course. Incidentally, some authors have missed the importance of these works. See *e.g.* a number of papers in G.S. Reynolds (ed.), *The Qur'ān in its Historical Context*, London, Routledge, 2008.

comparing a variety of Companion codices and picking the most common version where they differed. Stemmatics does not depend on the nature and form of the variants or the character of a codex as a whole.

The nature of the differences is taken up in my second approach to textual criticism, polarity analysis of internal evidence: I analyze two texts (the codices of C-1 and ʿUṯmān) to identify statistical asymmetries between them, and determine which text type is likely to be earlier. The inference from asymmetry to relative dating in this type of analysis follows principles that reflect the mode of transmission. The Companion codices demand principles that are suited to cases of concomitant written and oral transmission, as in dictation. The essay expounds such principles and applies them to the differences between the codices of ʿUṯmān and C-1. This approach, too, indicates that the wording of ʿUṯmān's codex is older than that of C-1.

The analysis paves the way for a discussion of the ancestor of all Qurʾānic textual traditions, the "prototype." The evidence analyzed in this essay indicates that the prototype is to be identified with the Prophet Muḥammad. Aside from the issue of dating, the evidence allows a discussion of other characteristics of the prototype. The conclusions here diverge from the usual pre-modern, traditional account towards a more conservative understanding. The most secure conclusion of the present study is that the sequences of verses and sentences were fixed already in the Prophetic prototype. This result is a departure from mainstream traditional views about the date when the pieces of revelation were joined together into fixed *sūra*s. It is traditionally more common to date this to the reign of ʿUṯmān. The findings here support the less common dating to the time of the Prophet.

Another way in which the conclusions here differ from traditional views concerns the relative merits of the different Companion codices. The majority traditional scholarly position associates Companion Qurʾāns with the Seven Modes (*aḥruf*) in which God revealed the holy book. It thus does not claim for the ʿUṯmānic Qurʾān greater verbal fidelity to the original than the codices of Companions such as Ibn Masʿūd or, by implication, C-1. (Nevertheless, the majority position considers ʿUṯmān's version as the only one fit for recitation in worship. This is justified by citing consensus or by calling it the Prophet's last version, *i.e.* Muḥammad's last presentation of the Qurʾān to, or from, Gabriel.)[4] On the other hand, an early minority strand of opin-

[4] On the majority pre-modern position on Companion codices, see Ibn al-Ǧazarī, *Taqrīb al-našr fī l-qirāʾāt al-ʿašr*, Cairo, Muṣṭafā l-Bābī, 1381/1961, p. 32-4; cf. Ibn ʿAṭiyya, ʿAbd al-Ḥaqq b. Ġālib b. ʿAbd al-Raḥmān, *al-Muḥarrar al-waǧīz fī tafsīr al-kitāb al-ʿazīz*, Beirut, Dār al-kutub al-ʿilmiyya, 1413/1993, p. 1-49. Ibn al-Ǧazarī writes that in his time important and unimportant scholars alike believed that seven modes (*aḥruf*) in which the Qurʾān was revealed

ion, espousing a kind of codical pluralism, considered non-ʿUṯmānic Qurʾāns as valid in all respects. Neither approach took the standard Qurʾān to be a more accurate representation of the Prophet's recitation. The present essay, however, finds that the textual evidence supports a more conservative conclusion, namely that the ʿUṯmānic tradition is better than C-1 in reproducing the Prophetic prototype. (This conclusion, however, may be modified in future if the parts of the manuscript not studied here do not support the trends observed thus far.)

When thinking about the text of an individual manuscript, it is necessary to understand it as a representative of a *text type*. The manuscripts belonging to the standard textual tradition, *i.e.* the ʿUṯmānic tradition, may have many differences among themselves; but their differences are nonetheless minor when these manuscripts are compared to the C-1 or Ibn Masʿūd texts. We may therefore speak of an ʿUṯmānic text type, of which each individual ʿUṯmānic manuscript is an instance despite the differences it may exhibit with other ʿUṯmānic texts. A text type is thus a cluster of similar texts which stand apart from other clusters. The text type is an essential notion for textual criticism, for behind it lies the concept of a textual tradition. Texts that are in the same text type are generally more closely related to one another in terms of origination. They are likely to belong to the same branch of the textual tradition. Text types thus allow one to arrange texts in a family tree just as similar phenotypes signify evolutionarily related species of animals and plants.

It will be convenient to refer to the manuscript studied here, *i.e.* to both of its layers, and to parts of it that are not extant, as Ṣanʿāʾ 1. The Companion of the lower codex, whoever he may have been, the codex of this Companion, *and* its corresponding text type and textual tradition, will be referred to as C-1, standing for "Companion 1". Context will disambiguate the sense. Similarly, the names ʿUṯmān and Ibn Masʿūd will be applied equally to the persons, the codices associated with them, and the resulting text types.

It should be noted that, throughout this essay, by "ʿUṯmānic", I do not mean "adhering to the ʿUṯmānic orthography", which is the sense in which some writers use this adjective. Rather, I mean "belonging to the text type of the standard Qurʾān". Standing behind this text type is the textual tradition of the standard Qurʾān—a tradition that reportedly began with the codices ʿUṯmān disseminated as the ancestors of all the manuscripts in the textual tradition. In the course of dozens or hundreds of copyings intervening since

---

went well beyond the ʿUṯmānic text. He quotes one authority to the effect that other modes can include additions, omissions, substitutions, and transpositions of words. These are precisely the types of variants found in C-1 and the reported codex of Ibn Masʿūd. One may add that some pre-modern *tafsīr*s quoted the codex of Ibn Masʿūd.

the time of the ancestor manuscripts, many small changes may have accumulated. By far, the most common type of change concerns spelling. The spellings of Arabic words were not standardized and an early scribe often followed his own discretion rather than the conventions of the manuscript he was copying. Thus, an ʿUṯmānic manuscript, in the sense I use the term, needs not preserve the original orthography of its ancestor manuscript. It may also differ from the ancestor manuscripts in ways that are more significant.

Ṣanʿāʾ 1 Folios Identified and the Results of Radiocarbon Dating Described

The research reported here began with the examination of a single folio of a palimpsest, called here "Stanford '07", of dimensions 36.3 cm by 28.5 cm. The study then extended to photographs of some of the other leaves from the same manuscript. A palimpsest is a parchment with at least two layers of text. The older layer is scraped off or washed to make room for the new text, but over time it resurfaces as a shadow, in this case as a pale brown text. In August 2007, Uwe Bergmann subjected the Stanford '07 folio to X-Ray fluorescence imaging at the Stanford Synchrotron Radiation Laboratory in the Stanford Linear Accelerator Center (SLAC). X-Ray fluorescence imaging is a technique for detecting, and tracing, the chemicals left on a leaf by inks or dyes. Its application to the Stanford '07 folio assisted in reading and tracing out the lower text, bringing to light some letters, verse separators, and diacritical marks not otherwise visible or legible (Figures 3-4). Moreover, the fact that the inks used in the two layers were chemically different made it possible to determine to which layer every feature belongs. For example, it was possible to confirm that the diacritical marks and verse dividers were in the same ink as the main text. This holds true for both the lower and upper texts. The same is true of the decorative *sūra* separators of the lower text. Therefore, these features probably were not added at a later stage. High-resolution photographs of the folio, images of the traces of chemicals, and a hand-drawn tracing of the lower text can be downloaded from http://ssrl.slac.stanford.edu/quranleaf.

Radiocarbon dating was performed on a sample of Stanford '07. The analysis was done at the Accelerator Mass Spectrometry (AMS) Laboratory at the University of Arizona. The results indicate that the parchment has a 68% probability of belonging to the period between AD 614 to AD 656. It has a 95% probability of belonging to the period between AD 578 and AD 669.[5]

[5] The results of radiocarbon dating at the NSF Arizona AMS Laboratory at the University of Arizona were described in a letter dated May 23, 2008, by the director of the lab and Professor of Geosciences and Physics, A.J. Timothy Jull. What follows is a quotation from the letter:

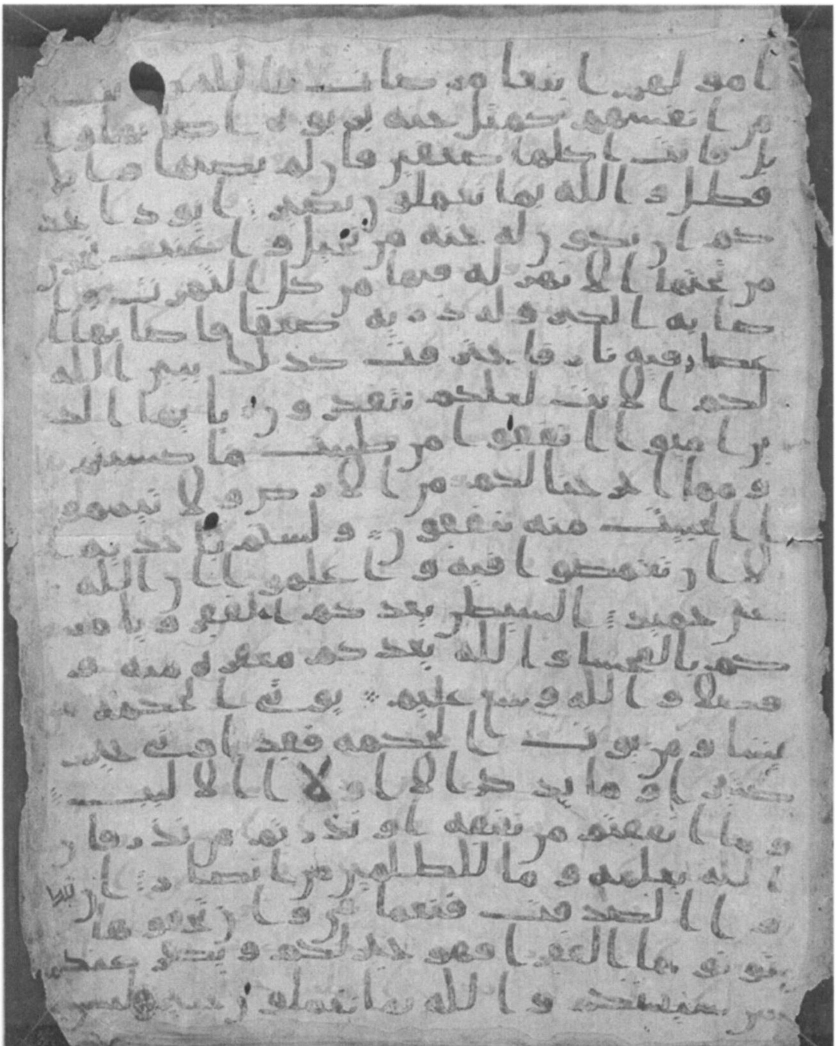


Figure 1. Stanford '07, recto.

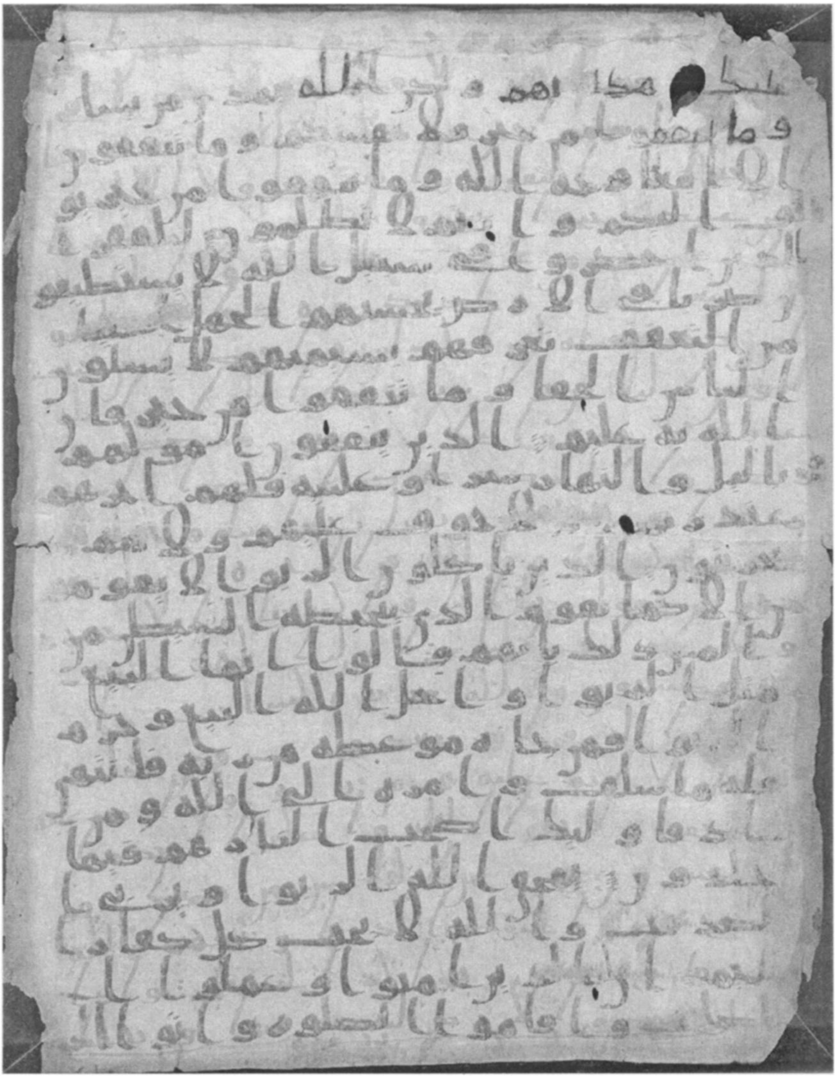


Figure 2. Stanford '07, verso.

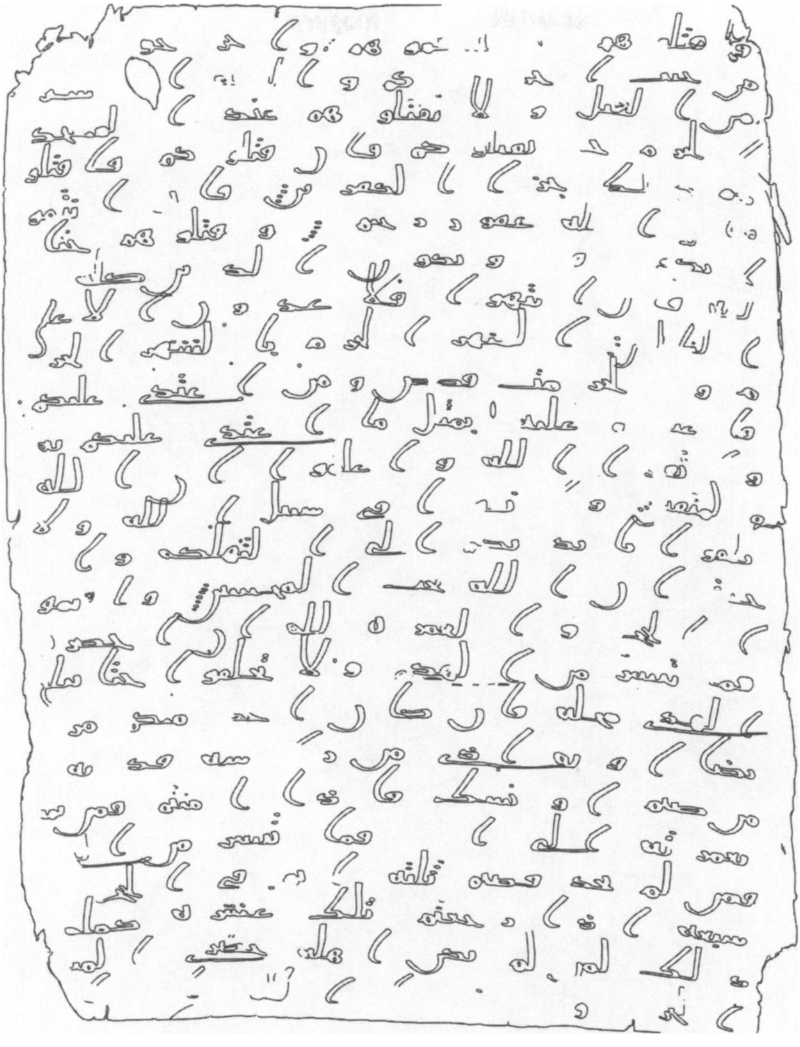


Figure 3. Stanford '07, lower text, recto. Kor 2, 191-6; or by the folio's own numbering, Kor 2, 188-93.

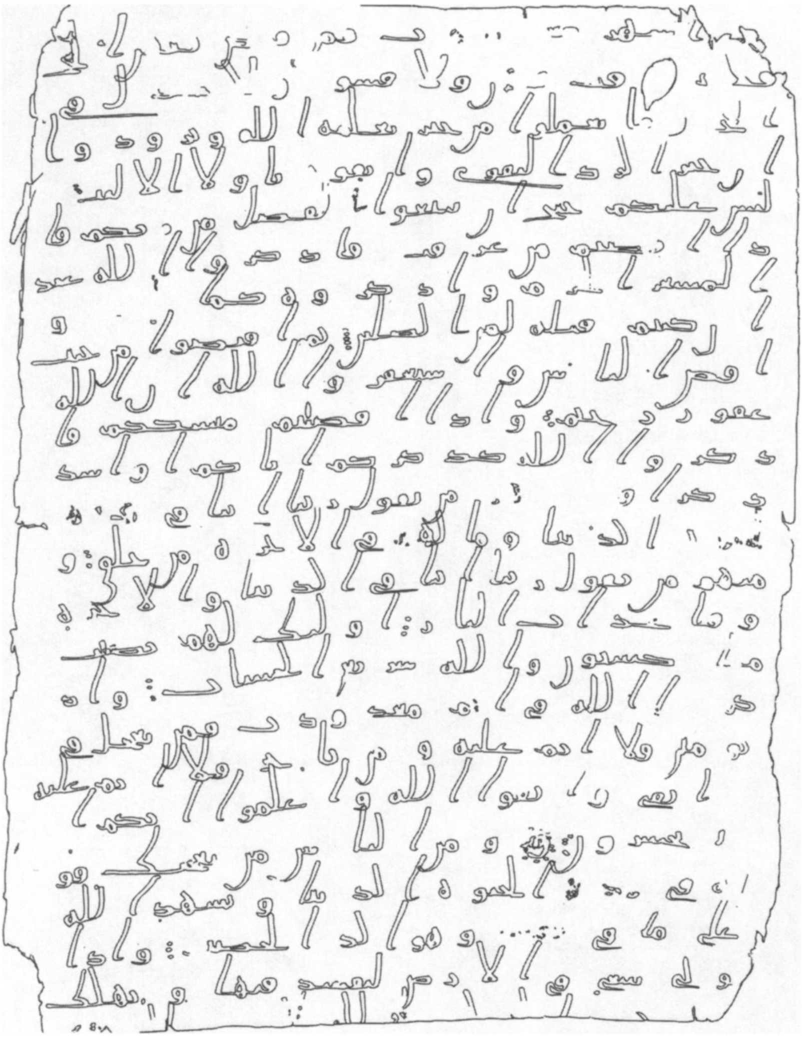
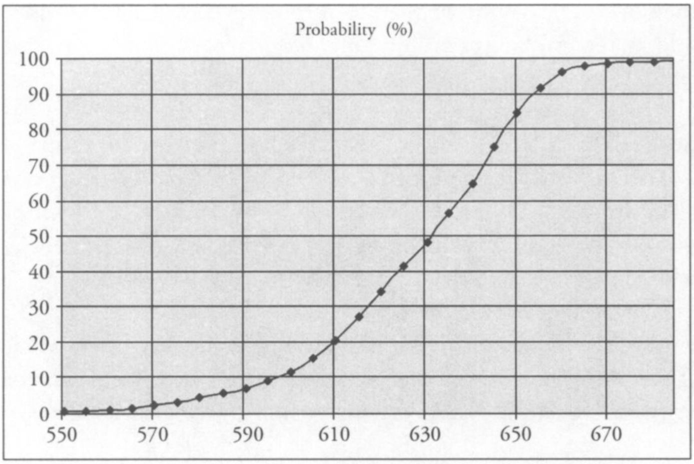


Figure 4. Stanford '07, lower text, verso. Kor 2, 197-205; or by the folio's own numbering, Kor 2, 194-202.

For historical reasons, however, what is of greater interest is the probability that the parchment is older than a certain date. Such probabilities are readily calculated and are shown in Table 1.[6] The probability that the parchment is older than AD 646 is 75.1 %, or a three-to-one likelihood. It is highly probable therefore, that the Ṣanʿā' 1 manuscript was produced no more than 15 years after the death of the Prophet Muḥammad.

Table 1. ***Result of Radiocarbon Dating. The graph gives, in five-year increments, the probability (%) that the parchment is older than a certain date AD, while the table to the right gives the probabilities for selected years.***



Date (AD)	Probability that the parchment is older
635.5	56.2 %
645.5	75.1 %
655.5	91.8 %
660.5	95.5 %
670.5	98.8 %
675.5	99.2 %

"$\delta^{13}C$	−20.8 per mil
Fraction modern	0.8393 ± 0.0037
Radiocarbon Age	1,407 ± 36 years BP
Calibrated Age	614-656 AD (1 sigma, 68% confidence)
	578-669 AD (2 sigma, 95% confidence)

The 14C measurements were done by accelerator mass spectroscopy (AMS) and I have also quoted the stable-isotope value of $^{13}C/^{12}C$, given in $\delta^{13}C$ units. The values are quoted corrected to −25% for $\delta^{13}C$, which indicates the value of stable-isotope ratio of $^{13}C/^{12}C$ deviation from a known standard, in parts per mil (‰). Most organic materials are typically about −25‰.

The radiocarbon age is the conventional ^{14}C age and is quoted in years 'before present,' where 'present' has been defined as the expected natural level for ~1950 AD. As the radiocarbon content of the atmosphere fluctuates with time, it is necessary to calibrate this value against known material, usually known-age tree rings."

[6] I obtained the probability distribution from the radiocarbon age 1,407 ± 36 years BP using the IntCal 04 calibration and the software on the website of the Oxford Radiocarbon

The date of the parchment is a reliable indicator of the date of the lower writing. The parchment probably is not many years older than the lower writing. Given its dimensions, this manuscript must have been expensive, requiring a whole flock of animals. It is unlikely that the folios required for this Qurʾān were procured for a purpose other than the one to which they were put. In the initial decades of Islam, the period to which this manuscript belongs, the Arabs did not have many books to copy beside the Qurʾān. Indeed, the only extant vellum manuscripts of a comparable size in the Ḥiǧāzī script are without exception Qurʾāns. There would not have been a large supply of unused folios of this size.

There is no public record of the Stanford '07 folio prior to its being auctioned by Sotheby's in London in 1993. At least three other folios from Ṣanʿāʾ 1 have been sold in auction houses in London, and images of these have been used in the present study. A leaf that would have been adjacent to Stanford '07 in the codex was auctioned by Sotheby's in 1992, resurfaced in a Christie's auction on May 1, 2001, emerged in the Sam Fogg Gallery,[7] and finally moved to the David Collection in Copenhagen, where it now resides. A leaf was auctioned by Bonhams in 2000, and another by Christie's on April 8, 2008.[8] [9] At least thirty-two other leaves from Ṣanʿāʾ 1 are in Ṣanʿāʾ, Yemen, in the collection of manuscripts retrieved from the Grand Mosque of Ṣanʿāʾ, bearing the catalog number 01-27.01.

Table 2 identifies the folios included in this study. The quality of the available images varies. For example, for the Bonhams folio I was limited to the image in the printed auction house catalog, while the David Collection kindly provided very high-resolution photographs. Of the leaves in Ṣanʿāʾ, a few low-resolution images have been published, useful only for the study of the upper text. My treatment of the lower text, therefore, is limited to the four folios sold in London.

The upper writing is the work of two scribes. One scribe wrote the first two folios (from the end of *sūra* 2), and another scribe wrote the others. Each

Accelerator Unit (https://c14.arch.ox.ac.uk/login/login.php?Location=/oxcal/OxCal.html). The results are consistent with the calibrated age announced by Timothy Jull. I subsequently calculated running totals to obtain the probabilities for one-sided intervals shown in Table 1.

[7] Marcus Fraser and Will Kwiatkowski, *Ink and Gold: Islamic Calligraphy*, London, Paul Holberton, 2006.

[8] The Stanford '07 folio was first disclosed in Sotheby's, *Oriental Manuscripts and Miniatures*, 1993, p. 18-23. For the other leaves see *e.g.* Sotheby's, *Oriental Manuscripts and Miniatures*, 1992, p. 254-9; Bonhams Knightsbridge, *Islamic and Indian Works of Art*, 2000, p. 11-4; Christie's, *Art of the Islamic and Indian Worlds*, 2008, p. 24-7.

[9] The Christie's folio was sold for $4,924,279, breaking the world auction record for any Islamic manuscript (http://www.christies.com/presscenter/pdf/04082008/144844.pdf, accessed in August 2008).

Table 2. ***The folios of Ṣanʿāʾ 1 studied here.***

Name of Leaf and Source	Upper Text	Lower Text
Stanford '07 (Originally Sotheby's '93)	Recto: 2, 265-271 Verso: 2, 271-277	Recto: 2, 191 (first word) to near the end of 2, 196. Verso: 2, 197 (*ašhurun*)—2, 205 (the last fully visible word is *yuhlika*, last inferred word *al-fasāda*).
David Collection (originally Sotheby's '92), Copenhagen, inv. no. 86/2003.	Recto: 2, 277-282 Verso: 282-286	Recto: 2, 206 (*wa-iḏā*)—2, 217 (the last visible word is *wa-l-masǧidi*, and the last inferred word *minhu*). Verso: 2, 217 (*akbaru ʿinda llāhi*)—2, 223 (*ḥarṯun lakum*).
Bonhams 2000	Recto: 4, 33-43 Verso: 4, 43-56	Recto: 5, 41 (*yuṭahhira*)—5, 47 (*fa-ulāʾika hum*). Verso: 5, 48 (... *li-mā bayna yadayhi min al-kitābi muhayminan*; this last is the first mostly visible word)—5, 54 (*llāhi wa-lā*).
Christie's 2008	Recto: 4, 171-5, 3 Verso: 5, 3-9	Recto: Beginning of 63 (*bi-smi llāhi l-raḥmāni l-raḥīmi*) to the end the *sūra*; the beginning of 62, with the *basmala*, to 62, 11 (*tiǧāratan aw lahwan*). Verso: 62, 11 (*infaḍḍū*) to the end of the *sūra*; entire *sūra* 89 with the *basmala*; *sūra* 90 (with the *basmala*) to 90, 6 (*yaqūlu*).
Noseda 1a UNESCO Image 152255B	6, 49—6, 61	The image available to the author is not legible.
Noseda 2a UNESCO Image 152254B	6, 149—6, 159	The image available to the author is not legible.

Table 2 (*cont.*)

Name of Leaf and Source	Upper Text	Lower Text
Noseda 4a UNESCO Image 027002B	16, 73-89	The image available to the author is not legible.
Noseda 5a UNESCO Image 027004B	16, 102-118	The image available to the author is not legible.
Noseda 8a UNESCO Image 152256B	20, 1—20, 43	The image available to the author is not legible.
Noseda 14b *Maṣāḥif Ṣanʿāʾ*, 1985, inside cover, partial image.	28, 74—28, 86	15, 54-72
Noseda 17a *Maṣāḥif Ṣanʿāʾ*, 1985	31, 24—32, 4	Possibly 9, 103-113. The image available to the author is not legible.
Noseda 18a Hans-Casper von Bothmer, 1987	32, 20—33, 6 (a bifolio)	The image available to the author is not legible.
Noseda 31a UNESCO Image 152257B	59, 2—59, 10	The image available to the author is not legible.

had a distinct handwriting and distinctive dotting and orthographical habits. For example, the first scribe tended to dot the *tā*'s and spelled *ʿalā* with *yāʾ* (على), whereas the second one dotted the *yā*'s and spelled *ʿalā* with *alif* (علا).[10]

[10] François Déroche has noted, "in manuscripts written by two or more copyists, it is not infrequent to notice that each of them had his own way of putting diacritical dots, with regard either to their density, or to the individual letters receiving the dots" (François Déroche, "New Evidence about Umayyad Bookhands", in *Essays in Honour of Salah al-Din al-Munajjid*, London, al-Furqan Islamic Heritage Foundation, 2002, p. 611-42, 627).

Table 3. ***The four hands in Ṣanʿāʾ 1, listed in the proposed chronological order. This essay is concerned exclusively with the first two.***

Name of Hand	Prevalence	Appearance	Iron & Copper in Ink
Lower Text	systematic	Light brown	Yes
Upper Text	systematic	Dark brown	Yes
Upper Modifier	occasional	Black, poor hand	No
Lower Modifier	occasional	Black, excellent hand	No

Both literary and paleographic evidence show that there was nothing unusual about two or more persons collaborating to copy a Qurʾān.[11] As for the lower text, it appears to be entirely in the hand of one scribe, although a closer examination is needed to verify this. Interestingly, the first two folios listed in Table 2 are adjacent in both the lower *and* upper layers.

Basic Description of the Folios

There are four layers of writing on Ṣanʿāʾ 1. These are shown in Table 3. The "lower text" and "upper text" apparently constituted complete codices of the Qurʾān. The upper one is ʿUṯmānic and the lower one is not. The "upper modifier" (in the David and Stanford '07 leaves) appears as an awkward, amateurish hand that wrote over the words where the upper text had faded, especially near the edges. The hand designated as "lower modifier" (in the David and Stanford '07 leaves) did two things. First, it wrote over some words and letters of the lower text that had faded or had not resurfaced because the ink had been thoroughly removed by the original scraping off of the lower writing.[12] Second, on occasion it wrote the standard version where

[11] For literary evidence confirming that there could be collaboration in copying, see the report of ʿAlī b. al-Ḥusayn b. ʿAlī (d. 94/712), of which there are at least two versions related through his grandson, Ǧaʿfar, as the common link: Abū Bakr ʿAbd Allāh b. Abī Dāwūd Sulaymān b. al-Ašʿaṯ al-Siǧistānī, *Kitāb al-Maṣāḥif*, ed. Arthur Jeffery, Egypt, Maṭbaʿat al-Raḥmāniyya, 1355/1936, p. 166; Abū Bakr Aḥmad b. al- Ḥusayn b. ʿAlī l-Bayhaqī, *al-Sunan al-kubrā*, Dār al-fikr, Beirut, n.d., VI, p. 16. The paleographic evidence consists of early manuscripts that are in more than one hand. Such manuscripts are invaluable for the paleographer, for they give evidence of the range of variation in calligraphic and orthographic practices in a single time and place. See François Déroche, "New Evidence about Umayyad Bookhands", p. 611-42, 629; *id.*, *La transmission écrite du Coran dans les débuts de l'islam: Le codex Parisino-petropolitanus*, Leiden, Brill, 2009, p. 26-77; Intisar Rabb, "Non-Canonical Readings of the Qurʾān: Recognition and authenticity (The Ḥimṣī reading)", *Journal of Qurʾānic Studies*, 8/2 (2006), p. 84-127.

[12] Here is an example of the lower modifier writing over a faded or damaged part of the

the lower text departed from it. However, it does neither of these things systematically. And it does not try to obscure the deviations from the standard version. The evidence shows that the lower modifier came after the erased lower text had reemerged, hence after the upper text.[13] It could even be an early modern intervention. So, the lower and upper texts are the earliest ones. The question arises which script came earlier: the upper-modifier or the lower-modifier. The upper-modifier seems to have been added to ensure the continued usability of the codex despite the faded words. The lower-modifier seems to have been added at a time when usability was no longer a concern. I conjecture, therefore, that it came last. A thorough physical examination is needed, however, before the chronological order of the last two writings is settled beyond doubt. In any case, among the four layers, the upper and lower texts are the oldest, while the "modifier" texts, having come later, hold relatively little interest. *It is thus with the lower text and upper text that this essay is exclusively concerned.*

A few noteworthy differences between the upper and lower writings may be mentioned. The upper script is more generous than the lower one with the tiny dashes that distinguish consonants of a similar shape. The fact that the lower script exhibits such marks is important; but it is not wholly unexpected, since a commercial papyrus from the year 22/642 also has such marks.[14] So does an inscription from the year 24/644-5,[15] in addition to other early documentary sources.[16] Moreover, pairs of such dashes are arranged, more

lower text: most of the *mīm* of *al-ḫiṣāmi* in Kor 2, 204 had been scraped off, and no trace of the ink is detectable by X-Ray Fluorescence Imaging. Nonetheless, we know that there was a *mīm* there because the telltale tip of its tail survived damage. The lower modifier wrote a full *mīm* in the expected place.

[13] The strongest indication that the lower modifier came after the upper text is that it was not erased before the upper text was written. Another indication that the lower modifier came after both the lower and upper writings is that its ink is chemically unlike theirs and similar to that of the upper modifier. It contains no iron or copper and is black, being apparently carbon-based. By contrast, the inks of the lower and upper writings both contain iron and copper. Finally, the calligraphic style of the lower modifier suggests that it did not belong to the first two centuries AH. The lower modifier gives one the impression of somebody who read the emerged lower text out of curiosity and casually noted some discrepancies with what he remembered of the Qur'ān.

[14] Alan Jones, "The Dotting of a Script and the Dating of an Era: The strange neglect of PERF 558", *Islamic Culture*, 72 (1998), p. 95-103.

[15] ʿAli ibn Ibrahim Ghabban, Robert Hoyland, "The inscription of Zuhayr, the oldest Islamic inscription (24 AH/AD 644-645), the rise of the Arabic script and the nature of the early Islamic state", *Arabian Archaelogy and Epigraphy*, 19/2 (2008), p. 210-37.

[16] There are even pre-Islamic inscriptions with diacritical marks distinguishing consonants. For a non-exhaustive list and images of some of the early sources with such marks see Muḥammad Muṣṭafā l-Aʿẓamī, *The History of the Qur'ānic Text*, 2nd ed., Riyadh, Azami Publishing House, 2008, p. 152-6.

often than not, with an upward slope in the upper script and a downward slope in the lower one. The upper writing has more words per page than the lower one. The upper layer uses a special marker for every ten verses. By contrast, the lower text marks the 200th verse of *sūra* 2 by means of what looks like the letter *wāw* encircled by short dashes. The symbol is similar to those used in the BNF arabe 328 (a) manuscript to mark groups of fifty verses (*i.e.* the 50th, 100th, and 150th verses).[17] The chemical difference between the inks used has already been noted.

There are a number of ways in which the lower text is more "elaborate" than the upper one, thus complicating the assumption that more sophisticated means later. The lower text has decorations between *sūra*s, but the upper text does not.[18] The lower text inserts captions between *sūra*s such as, "This is the end of *sūrat al-Munāfiqīn*"; but the upper text lacks *sūra* names. There is resonance here with literary sources: certain first-century figures such as Ibrāhim al-Naḫaʿī (d. 96/715, Kūfa) and ʿAṭāʾ (d. 114/732, Mecca) were opposed to captions that said, "This is the end of *sūra* such and such", or stated the number of verses in the *sūra*.[19]

Surprisingly, the lower script on occasion appears to use what are possibly diacritics, in the form of perfectly round dots, to signify short vowel marks (and possibly elided *alif*s, *i.e. hamzat al-waṣl*). These dots are in the same ink as the rest of the lower writing and do not appear to have been added later. In particular, the words *tahluka* and *nusukin* in Stanford '07, in Kor 2, 195 and 196, appear to have dots on the consonants *lām* ("l") and *sīn* ("s") respectively, both placed at an elevated level, possibly for indicating the vowel "*u*" (the *ḍamma*). The vocalization of both of these letters later became a matter of dispute. So, it seems unlikely that the dots are in these positions by chance. If these words are vowelled, this may indicate that they were already deemed difficult at the time of C-1. In each case, the lower writing's apparent vowelling agrees with what came to be the majority view of later authorities.[20]

---

[17] See François Déroche and Sergio Noseda, *Sources de la transmission manuscrite du texte coranique: I Les manuscrits de style ḥiǧāzī, Volume i, Le manuscrit arabe 328 (a) de la Bibliothèque nationale de France*, Lesa, Fondazione Ferni Noja Noseda, 1998, p. 29 (Kor 3, 100), p. 107 (Kor 6, 50), p. 116 (Kor 6, 100), p. 155 (Kor 7, 150).

[18] The fact that the lower text sometimes has decorative markings between *sūra*s has already been noted by Sheila Blair, *Islamic Calligraphy*, Edinburgh, Edinburgh University Press, 2006, p. 119, with regard to one of the folios located in Ṣanʿāʾ, namely Noseda 18a. Such *sūra* dividers are found also in the Christie's folio, but not some other folios.

[19] Ibn Abī Dāwūd, *al-Maṣāḥif*, p. 137-41.

[20] For the disagreements over the vowelling of these words, see ʿAbd al-Laṭīf Muḥammad Ḫaṭīb, *Muʿǧam al-qirāʾāt*, Damascus, Dār Saʿd al-Dīn, 2002, I, p. 266 and 269.

However, more folios need to be analyzed before one states with certitude that the dots are indeed diacritical marks, as opposed to accidental ink droplets.

Apart from these distinctions, the upper and lower scripts have fairly similar calligraphic styles.[21] They also use a similar end-of-verse marker in the form of several short dashes; but in the upper layer these are often arranged in triangles. Both writings lack additional spacing between words; a word may be split at the end of the line; the number of lines can vary from one page to the next; and the lines of text are not always strictly straight or perfectly horizontal.[22] Another shared feature of the two layers is the distinctive spelling اولا الالبب (in the accusative) instead of اولى الالبب. This spelling is attested in several other Qur'āns in the Ḥiǧāzī script.[23]

The non-ʿUṯmānic status of the lower text is shown by two things. First, its verbal differences with the ʿUṯmānic Qur'ān show that it represents a different text type and may therefore belong to a different branch of the family tree. Moreover, its variants are similar *in nature* to those reported of Companion codices. The second indication is its *sūra* ordering. On the Christie's folio, *sūras* 63, 62, 89, and 90 appear in that order, a sequence that is very close to what has been reported for the codex of Ubayy b. Kaʿb, a point that will be discussed further (p. 51). In terms of wording, the lower text also agrees with reported non-ʿUṯmānic variants in a few cases, as shown in Table 4; however, as a rule, reported non-ʿUṯmānic variants do not appear in C-1, nor are the variants of C-1 reported in the sources. Thus C-1 should not be identified with the codices whose variants have been described in the literary sources (Ibn Masʿūd or Ubayy b. Kaʿb); it represents an independent codex, text type, and textual tradition.

---

[21] Based on a preliminary examination of the lower and upper scripts, there appears to be no letter shape in one script that is not attested in the other, although the two scripts may use the different shapes of the same letter with different frequencies. In the upper writing, the letters are more angular; for example, the *nūns* sometimes display sharp turns.

[22] This last point, the lack of linear discipline, has been noted already in Sotheby's, *Oriental Manuscripts and Miniatures*, 1993, p. 18-23, 22.

[23] In some other early Qur'āns, the term *ūlū l-albābi* is spelled with an *alif* in the nominative and accusative cases and with a *yā'* in the genitive. See *e.g. Maṣāḥif Ṣanʿā'*, Dār al-āṯār al-islāmiyya, Kuwait National Museum, 1405/1985, no. 11 (catalog number 00-29.1). The MS arabe 328 (a) in the Bibliothèque nationale de France has an *alif* in Kor 3, 7 (p. 18), Kor 13, 19 (p. 213), and Kor 14, 52 (p. 225). It has a *yā'* in Kor 3, 190 (p. 6) and Kor 12, 111 (p. 209). The British Library's Or. 2165 has an *alif* in Kor 13, 19 (p. 123) and Kor 14, 52 (p. 136). It has a *yā'* in Kor 12, 111 (p. 119). See Déroche, *Sources... arabe 328(a)*; *id., Sources de la transmission manuscrite du texte coranique: I Les manuscrits de style ḥiǧāzī, Volume ii, Le manuscrit Or. 2165 (f. 1 à 61) de la British Library*, Lesa, Fondazione Ferni Noja Noseda, 2001.

Table 4. *Instances where C-1 agrees with a reported non-ʿUṯmānic variant. Words that are only partially legible are in parantheses.*[24]

Verse	ʿUṯmān	C-1	Codex agreeing with C-1
2.217	*qitālin fīhi*	*(wa)-ʿan qitālin fīhi*	*ʿan qitālin fīhi*
	"They ask you about the holy month—fighting in it."[25]	"They ask you about the holy month and about fighting in it."	"They ask you about the holy month, about fighting in it." (Reading of Ibn Masʿūd.)[26]
2.222	*fa-ʿtazilū l-nisāʾa fī l-maḥīḍi wa-lā taqrabūhunna ḥattā yaṭhurna*	*(fa-lā taqrabū) l-nisāʾa fī (maḥīḍihinna) ḥattā yataṭahharna*	*wa-lā taqrabū l-nisāʾa fī maḥīḍihinna wa-ʿtazilūhunna ḥattā yataṭahharna*
	"Avoid sexual relations with women during menstruation and do not go to them until they are cleansed"	"Do not go to women during their menstruation until they are clean"	"Do not go to women during their menstruation and avoid sexual relations with them until they are clean" (The reading of Ibn Masʿūd and the Codex of Anas. Both Ubayy b. Kaʿb & Ibn Masʿūd preferred *yataṭahharna*.)[27]

[24] The first three cases in the table have already been mentioned by Alba Fedeli, "Early Evidences of Variant Readings in Qurʾānic Manuscripts", in *Die dunklen Anfänge: Neue Forschungen zur Entstehung und frühen Geschichte des Islam*, K.-H. Ohlig & G. Puin (eds), Berlin, Verlag Hans Schiler, 2007, p. 293-316.

[25] My translations of Qurʾānic passages are often based on those of other translators. I have consulted the following: M.A.S. Abdel Haleem, *The Qurʾān: A New Translation*, Oxford, Oxford University Press, 2004; Arthur Arberry, *The Koran Interpreted*, London, Oxford University Press, 1983; Muhammad Asad, *The Message of the Qurʾān*, Gibraltar-London, Dar al-Andalus, 1980; Abdalhaqq Bewley, Aisha Bewley, *The Noble Qurʾān: A New Rendering of its Meaning in English*, Norwich, Bookwork, 1999; Thomas Cleary, *The Qurʾān: A New Translation*, Chicago, Starlatch Press, 2004; E.H. Palmer, *The Qurʾān: Translated by E.H. Palmer*, Oxford, The Clarendon Press, 1900; Marmaduke Pickthall, *The Meaning of the Glorious Qurʾān: Text and Explanatory Translation*, Beirut, Dār al-kitāb al-lubnānī, 1971; J.M. Rodwell, *The Koran*, London, Dent, 1978; George Sale, *The Koran*, New York, Garland, 1984; Muḥammad Saʿīd Šakīr, *The Glorious Qurʾān*, Qum, Anṣāriyān, 1998; Mawlawi Sher Ali, *The Holy Qurʾān: Arabic Text and English Translation*, 14th ed., Surrey, Islam International, 1989; ʿAbdullāh Yūsuf ʿAlī, *The Meaning of the Holy Qurʾān*, Brentwood, Md., Amana, 1991.

[26] Ḫaṭīb, *Muʿǧam al-qirāʾāt*, I, p. 298.

[27] Ḫaṭīb, *Muʿǧam al-qirāʾāt*, I, p. 308-9.

Table 4 (*cont.*)

Verse	ʿUṯmān	C-1	Codex agreeing with C-1
5, 45	*wa-katabnā ʿalayhim*	*wa-katabnā ʿalā banī Isrāʾīla*	Either *wa-anzala llāhu ʿalā banī Isrāʾīla* or *wa-anzalnā ʿalā banī Isrāʾīla.*
	"We prescribed for them"	"We prescribed for the Children of Israel"	Either "God sent down upon the Children of Israel" or "We sent down upon the Children of Israel" (Codex of Ubayy b. Kaʿb)[28]
63, 7	*yanfaḍḍū*	*yanfaḍḍū min ḥawlihi*	The quotation of the verse in a *ḥadīṯ* is identical to C-1.
	"that they disperse"	"that they disperse from around him"	

The last item in Table 4 (*viz.* Kor 63, 7) is fascinating because here C-1 matches an unidentified codex that is quoted in passing in a *ḥadīṯ*, a tradition of Zayd b. Arqam. Five different transmitters related the tradition through this chain of transmitters: Isrāʾīl [b. Yūnus b. Abī Isḥāq] (d. AH 160-2)—Abū Isḥāq [al-Sabīʿī, ʿAmr b. ʿAbd Allāh] (d. 127, Kūfa)—Zayd b. Arqam (d. 65-8, Kūfa)—Prophet. In this report, Zayd discusses the circumstances surrounding the Prophet's dissemination of Kor 63, 7, and in doing so he quotes the verse. His quotation contains the phrase *min ḥawlihi*, which is found in C-1 but not in the standard Qurʾān.[29] This means that one of the persons in the above chain of transmission knew the verse with this extra element.

[28] Ḫaṭīb, *Muʿǧam al-qirāʾāt*, II, p. 279.

[29] Ibn Ḥanbal, Abū ʿAbd Allāh Aḥmad b. Muḥammad, *al-Musnad*, Lebanon, Dār Ṣādir, IV, p. 373; al-Buḫārī, Abū ʿAbd Allāh Muḥammad b. Ismāʿīl, *Ṣaḥīḥ al-Buḫārī*, n.p., Dār al-fikr, 1401/1981, VI, p. 63-5; Abū l-Ḥusayn Muslim b. Ḥaǧǧāǧ, *Ṣaḥīḥ Muslim*, Beirut, Dār al-fikr, VIII, p. 119; al-Tirmiḏī, Abū ʿĪsā Muḥammad b. ʿĪsā, *Sunan al-Tirmiḏī*, 2nd ed., n.p., Dār al-fikr, 1403/1983, V, p. 88; al-Nasāʾī, Aḥmad b. ʿAlī b. Šuʿayb, *al-Sunan al-kubrā*, Beirut, Dār al-kutub al-ʿilmiyya, 1411/1991, VI, p. 492; al-Ḥākim al-Nīsābūrī, *al-Mustadrak ʿalā l-Ṣaḥīḥayn*, Dār al-maʿrifa, Beirut, II, p. 488; al-Ṭabarānī, Sulaymān b. Aḥmad b. Ayyūb Abū l-Qāsim, *al-Muʿǧam al-kabīr*, 2nd ed., ed. Ḥamdī ʿAbd al-Maǧīd, n.p., Dār iḥyāʾ al-turāṯ al-ʿarabī, V, p. 186, 189; cf. Ibn Ḥaǧar al-ʿAsqalānī, Aḥmad b. ʿAlī, *Fatḥ al-Bārī*, 2nd edition, Beirut, Dār al-maʿrifa, VIII, p. 494, where he rejects the view of those who inferred from this *ḥadīṯ* that the phrase *min ḥawlihi* had been part of a codex.

The non-ʿUṯmānic status of the lower text has been noted already by Sergio Noseda, Yasin Dutton, and Alba Fedeli.[30] In 2007, Fedeli published a report on two of the folios (David and Bonhams). She mentions a number of important variants and points out the first three items in Table 4. Aside from that, her discussion is slightly imprecise.[31] Moreover, she misreads the text at one point, mistakenly inferring an omission in the Bonhams folio, and then uses that to argue that the version in the ʿUṯmānic Qurʾān resulted from deliberate addition motivated by religio-political concerns.[32]

The upper text is definitely ʿUṯmānic. It exhibits the kinds and magnitude of deviation from the standard text that typically accumulated in the course of written transmission. It is to the ʿUṯmānic textual tradition and the place of the upper text in it that I now turn.

---

[30] Sergio Noseda, "La Mia Visita a Sanaa e il Corano Palinsesto", *Istituto Lombardo (Rendiconti Lett.)*, 137 (2003), p. 43-60; Anonymous, "'The Qur'an: Text, interpretation and translation' 3rd Biannual SOAS Conference, October 16-17, 2003", *Journal of Qurʾānic Studies*, 6/1 (2003), p. 143-5 (mentioning Dutton's paper, "Three Possibly pre-ʿUthmānic Folios of the Qurʾān"); Fedeli, "Early Evidences", p. 293-316.

[31] At one point Alba Fedeli refers to "the so-called text of ʿUṯmān, which has now become 'the standard text' since 1924" ("Early Evidences", p. 306). Since ʿUṯmān's text became official 1268 years before that date, she has something else in mind, namely the Egyptian edition. On another page (p. 315), she says that it was in the fourth/tenth century that the ʿUṯmānic text became official; but this is wrong. Here, the author appears to conflate Ibn Muǧāhid's canonization of *particular* ʿUṯmānic readings (against other equally ʿUṯmānic readings alongside non-ʿUṯmānic ones) with the canonization of the ʿUṯmānic codex (against non-ʿUṯmānic ones). For some other issues, see next footnote.

[32] Fedeli thinks that the phrase *ʿan dīnihi* in Kor 2, 217 is missing from the lower text, and this becomes the basis for her theory that the term was added *deliberately* to the ʿUṯmānic version. That is incorrect. First, she has missed the unmistakable lowermost part of the ***nūn*** of *ʿan* which has survived the damage to the parchment. So,*ʿan dīnihi* was part of the text after all. (*ʿan dīnihi* is present also in the Bonhams 2000 folio in Kor 5, 54.) Second, the entire sentence *wa-man yartadid minkum ʿan dīnihi fa-yamut wa-huwa kāfirun* is generally illegible due to damage to the parchment. It is, therefore, not clear how she is able to conclude that the words *ʿan dīnihi* in this verse are missing. Presumably the fact that the lower-modifier wrote *ʿan dīnihi* leads her to think that the words were not there originally; but, as shown above in footnote 12, the lower modifier sometimes wrote words that filled the gaps created by irremediable erasure. It is thus entirely possible that *ʿan dīnihi* was part of the lower text, was erased irremediably, and then was written again by the lower modifier. Third, even if the term were missing here, there would be no reason for considering the ʿUṯmānic wording as the later one, as opposed to the earlier one. Her choice in this regard and her assumption of deliberate change betray, perhaps, a slight measure of conspiracy-mindedness. Fourth, the scenario Fedeli constructs to explain what may have motivated the addition of *ʿan dīnihi* is unclear as it stands (see "Early Evidences", p. 314).

II. The ʿUṯmānic Textual Tradition and the Upper Writing of Ṣanʿāʾ 1

What We Know about the Codex of ʿUṯmān

A large number of Qurʾān manuscripts from the first and early second century AH are extant. These are dispersed around the world in libraries, museums, mosques, and private hands. With the exception of the lower writing of Ṣanʿāʾ 1, they all belong to the standard textual tradition, called ʿUṯmānic in accordance with the traditional account of its origin, which traces it to the recension promulgated by ʿUṯmān b. ʿAffān, the Companion of the Prophet Muḥammad who ruled the Muslim empire as the third caliph during 23-35/644-56. Sometime around AD 650, ʿUṯmān is said to have sent copies of the Qurʾān to Kūfa, Baṣra, and Syria and to have kept a copy with him in Medina.[33]

How much of the traditional account of ʿUṯmān's dissemination of the standard Qurʾān is true? The reports that accumulated in the literary sources about the early history of the text are not without problems. They contain apparent contradictions and gaps in certain aspects of what transpired *before* ʿUṯmān disseminated the text. However, there are specific elements relating to the act of dissemination itself that can be verified. We know at least two things. First, there is no doubt that it was ʿUṯmān who established the standard by sending out master copies to different cities. Second, there is strong reason to believe that the text he established was what the sources say it was as far as the skeletal-morphemic aspect is concerned (*i.e.* not including differences in spelling conventions).

Our knowledge of these fundamentals does *not* depend on uncritical acceptance of reports of individuals (*āṯār*). For the first piece of our knowledge, that it was ʿUṯmān who established the standard version, one may note two arguments, one inspired by a passing observation of Hossein Modarressi about the collective memory of the early communities, and the other inspired by recent studies of Michael Cook and others that have shown the reliability of the *qirāʾāt* literature. For the second proposition, that the skeletal-morphemic ʿUṯmānic text was what the sources say it was, I refer to Cook's findings. To these arguments may be added the circumstantial corroboration to be offered by the present essay: textual analysis indicates that the ʿUṯmānic text preserves the Prophetic prototype better than C-1, and yet the lower writing is older than AD 656, probably belonging to the fifteen years after the

[33] For the date of this event see above, p. 2, as well as footnote 2.

Prophet's death. Therefore, the 'Uṯmānic wording is older than that of another codex already from the era of the Companions.[34]

How can one know that it was 'Uṯmān who sent the regional codices? The first argument for this concerns collective memories. In the late first century and early second century AH, communities remembered having received the standard text from 'Uṯmān.[35] It was not only the different cities that had the same recollection, but also the different and at times clashing religious communities, including the proto-Šī'īs and the Ḫāriǧīs. No one traced the standard version to a source other than 'Uṯmān. The 'Alīds did not ascribe it to 'Alī. Even those who transmitted the variants found in non-standard codices (such as the codex of Ibn Mas'ūd) did not dispute the 'Uṯmānic origin of the standard version.[36] Reports of individuals were inconsistent on the

[34] Note that my argument about 'Uṯmān does not rely on reports of individuals. For the purpose of this essay, I generally reserve judgment on the accuracy of reports of individuals unless a report receives corroboration from other types of evidence. Nonetheless, the number of traditions that bear on the origins of the 'Uṯmānic recension directly or indirectly is very large, and each report deserves detailed analysis. For a study of just one tradition, see Harald Motzki, "The Collection of the Qur'ān: A Reconsideration of Western Views in Light of Recent Methodological Developments", *Der Islam*, 78 (2001), p. 1-34.

[35] My argument from the collective memory of early communities is inspired by Hossein Modarressi, "Early Debates on the Integrity of the Qur'ān: A Brief Survey", *Studia Islamica*, 77 (1993), p. 13-4. A variant of the argument was subsequently used by Fred Donner, although not specifically with respect to 'Uṯmān's standardizaton of the Qur'ān (Fred Donner, *Narratives of Islamic Origins*, Princeton, NJ, Darwin Press, 1998, p. 26-8); cf. Hugh Kennedy, *The Prophet and the Age of the Caliphates*, New York, Longman, 1986, p. 357.

[36] A report in Ibn Šabba l-Numayrī, *Ta'rīḫ al-Madīna*, Qum, Dār al-fikr, I, p. 7-8, states that the Umayyad governor of Kūfa, al-Ḥaǧǧāǧ (d. 95/714), had a codex copied and sent to the main cities. The report adds that, later, the 'Abbāsid caliph al-Mahdī also sent a codex to Medina. The report explicitly acknowledges the prior codex of 'Uṯmān. Al-Ḥaǧǧāǧ's codex belonged to the 'Uṯmānic textual tradition. This is indicated by a Baṣran report listing the eleven alleged "changes" al-Ḥaǧǧāǧ made to the Qur'ān. Nine of the eleven variants involve just single characters (Ibn Abī Dāwūd, *al-Maṣāḥif*, p. 49-50). The differences are well within the range of variations one expects to emerge naturally within a textual tradition. A close study of the variants shows that the Baṣran author of the report had simply assumed that one particular Baṣran copy belonging to the standard text type represented the original text sent out by 'Uṯmān. He thus naively assumed that the eleven differences with al-Ḥaǧǧāǧ's codex represented changes made by the despised governor. In any case, there is no chance that al-Ḥaǧǧāǧ could have dislodged the various regional branches of the 'Uṯmānic textual tradition especially outside Iraq, and there is no evidence that he attempted to do so either in or outside Iraq; but there is evidence that if he did try, he failed. Muslim reports about his use of coercion concern Ibn Mas'ūd's text type. Furthermore, nobody ever confused his codex with the original copy of 'Uṯmān. One should also dismiss a minority Šī'ī claim that God's explicit references to 'Alī were removed from the Qur'ān, as well as a similar early 'Abbāsid claim (reported by 'Abd al-Masīḥ al-Kindī) about God's explicit references to the 'Abbāsid and Umayyad dynasties. I intend to discuss all of the above matters in detail in a separate article devoted to the role of al-Ḥaǧǧāǧ.

history of the text *before* ʿUṯmān's promulgation of the standard version, but no one doubted that it was ʿUṯmān who disseminated it.

It is not possible to envision an Umayyad caliph establishing the standard version. To assume so would require explaining not only how the Umayyads managed to erase their empire-wide intervention from the memories of their subjects, whose statements have reached us in large numbers, in various cities, tribes, and sects, but also how it was possible to induce the same false memory about ʿUṯmān among all, including the dissidents who preferred ʿAlī to ʿUṯmān.

To be sure, entire populations can believe in falsehoods. But there are facts of a public nature that one expects people to know. For example, populations know at any given time who their king or caliph is. The standardization of the Qurʾān is something that people in the main cities would have known about. Imagine a city receiving a codex of the Qurʾān from the caliph, one intended to replace the codices in the hands of its dwellers. This would have been a very public event, an occurrence of which the whole society would have been aware, a sensitive affair that people would have talked about. The city would know under which caliph it got its Qurʾān. Now give free rein to imagination: suppose it was the caliph ʿAbd al-Malik who standardized the Qurʾān. The question is: how was he able to persuade everybody that they did not receive the codex when they thought they got it? How could he convince even the communities teeming with rebels, or those who did not care much for ʿUṯmān?

If we suppose that the recollection of the different cities, sects, and tribes changed over time, replacing the name of ʿAbd al-Malik with ʿUṯmān, then it becomes impossible to explain why the memories of all cities, sects, tribes, and individuals changed in exactly the same way. Relatively decentralized, and continually facing rebellion and dissension, the Umayyads were hard-pressed to preserve their political authority over their domain. Damascus did not attempt to micromanage religious opinion across the empire, nor would it have succeeded in creating a consensus to replace an earlier consensus if it had tried to. Even in the much more centralized times that were to follow under the ʿAbbāsid caliph al-Maʾmūn, the attempt to fix theological dogma would prove futile. Revisionist back-projection is a common thing in both pre-modern and contemporary understandings of history. But such a back-projection would have failed to change *without any trace of dissent* what would have been *common knowledge* in several cities that were home to numerous clashing or rebellious communities. This observation becomes even more salient when one notes that on the vast majority of legal questions, each city had a distinct profile. The routine disagreements among the cities shows a large measure of independent religious development during the first century of Islam.

While the above argument concerns collective recollections of a momentous occurrence, my second argument is about individuals' memories of minutiae, memories that recently have been shown to be accurate. The communities had more than just a recollection of the source of their Qur'āns. There were established scholarly traditions of reading them that varied from scholar to scholar and from region to region. Scholars began to document these traditions by compiling and classifying the differences in pronunciation and orthography, and variations in the way *sūra*s were divided into verses (*viz.* the locations of the verse separators). Their efforts culminated in the genre of *qirā'āt*, the so-called "Readings" literature. The extant books in this field were written much later than the periods they describe, relying on reports that purportedly had been handed down for several generations. Yet recent research, to be discussed below, has shown that when it comes to the minutiae of textual variations, the literature carries substantially accurate information about the first- and early second-century phase. (To be sure, none of this implies that the genre is immune to inaccuracies and errors of transmission or that it captured *everything* about the early period.)

If the *qirā'āt* literature is substantially credible with regard to minutiae, that is all the more reason to accept the historicity of the broad framework it takes for granted. To give a whimsical yet apt analogy, imagine a person who alleges that there is an eight-hundred-pound gorilla in a room. It is unreasonable to say that his claims are substantially right about fine features of the primate—the color of its fur, the sizes of its toes, etc.—but wrong about its very existence. A dogmatic skeptic could devise a scenario where this could happen, but it would not be the natural conclusion. By the same token, the most natural way to explain why the *qirā'āt* genre is accurate in many of its particulars is to accept the truth of its core assumption that the standard textual tradition goes back to 'Uṯmān. For instance, it is hard to imagine that people could remember accurately the archetypal regional codices *at the level of individual words*, as Cook has shown, but get wrong the name of the caliph who disseminated those codices. It would be as if tradition preserved accurately all the minutiae of the Battle of the Bulge from WWII but projected the event back to WWI.

Among recent studies that give credibility to the *qirā'āt* genre, two are particularly important for the purposes of my study: those of Cook and Yasin Dutton. Cook analyzed the thirty-six reported differences in the codices 'Uṯmān sent to the cities—the original codices, as opposed to their less pristine offspring.[37]

[37] Michael Cook, "The Stemma of the Regional Codices of the Koran", *Graeco-Arabica*, 9-10 (2004), p. 89-104.

These variants must have been introduced in Medina as the texts were copied off one another before being dispatched to the cities. They consist of small changes to the skeletal text that actually make a difference in pronunciation, usually changing one word to another (hence my label, "skeletal-morphemic" changes), as opposed to changing merely the spelling of a word. Typically, the meaning does not change. For example, in Kor 7, 75 the codex of Syria has *wa-qāla* وقال ("and he said") whereas others have *qāla* قال ("he said"). In a few cases, the meaning is affected slightly. For example, the codex of Syria has *wa-yanšurukum* وينشركم ("makes you spread") in Kor 10, 22 whereas others have the graphically similar term *wa-yusayyirukum* ويسيركم ("makes you journey").

Cook's investigation of the list of skeletal-morphemic variations which, according to early scholars, distinguished the ʿUṯmānic codices, allowed him to reconstruct the genetic relationships between and among the codices. The variations among the cities follow strikingly regular patterns. For example, the codex of Medina has no variant that is unique to it. Some variants belong uniquely to it and Syria, and others uniquely to it and Baṣra. But it never agrees with Kūfa against Syria or Baṣra.[38] When such patterns are taken into consideration, the upshot is that of the well over 100 conceivable stemmata that could describe which codex was copied from which, only several closely-related ones fit the data, and among these some are more probable than others.[39] On the one hand, the great regularity of the patterns requires an explanation other than chance, and, on the other hand, it is difficult to explain why the texts fit this very specific set of stemmata if not as a result of having emerged in accordance with one of them. Cook argues convincingly against the possibility that the descriptions of these codices were forged by pointing out that the variants of four cities show no sign of cross-contamination. To create the appearance of non-contamination would have required skills that the scholars did not have, namely knowledge of the logic of stemmata.[40] From this, Cook infers that "we have to do with genuine transmissions from an archetype",[41] thus showing the correctness of the reports about the regional codices. The historicity of the received data, in turn, is "a testimony to the continuing accuracy of the transmission of the variants"[42] in the *qirāʾāt* literature. (Elsewhere, he reaches a similarly positive conclusion about

[38] *Ibid.*, p. 92-3.
[39] *Ibid.*, p. 103.
[40] *Ibid.*, p. 103-4.
[41] *Ibid.*, p. 104.
[42] *Ibid.*, p. 103.

the accuracy of the literary sources in describing the personal codex of a contemporary of ʿUṯmān.)[43]

Whereas Cook's corroboration of the *qirāʾāt* variants turns on the analysis of their patterns of agreement and disagreement, Dutton and a number of other scholars compare the literature with external evidence in the form of manuscripts.[44] For example, the variants of the ʿUṯmānic Qurʾān in the British Library, which is among the earliest set of manuscripts, point precisely and redundantly to the Reading tradition ascribed to Ḥimṣ, a town located north of Damascus.[45] The match between the text of the ancient folios and the reported attributes of the Ḥimṣī reading in the *qirāʾāt* genre is improbable *a priori* in the way it is improbable for two randomly picked pieces of a jigsaw puzzle to fit. The match provides the kind of mutual corroboration akin to that of the conjoined pieces of the puzzle, in that each piece sheds light on the other: the manuscript tells us that the reported Ḥimṣī reading found in the *qirāʾāt* literature preserves genuine information about the earliest class of manuscripts, and the *qirāʾāt* tell us that the folios were the work of a scribe following a textual tradition associated with the town of Ḥimṣ. Likewise, as shown by Tayyar Altıkulaç, *et. al.*, a nearly complete early codex in Istanbul reveals a match with the *qirāʾāt*-reported skeletal-morphemic patterns of Medina, while the skeletal-morphemic text of the well-known Tashkent

[43] Michael Cook and M.M. al-Aʿẓamī discuss a copy of the Qurʾān belonging to the Medinese Mālik b. Abī ʿĀmir al-Aṣbaḥī (d. 74/693). His grandson, the famous jurist Mālik b. Anas (94-179/712-795) showed it to his students, who recognized that its text matched the *muṣḥaf* of Medina up to *sūra* 40 and then switched to the *muṣḥaf* of Baṣra. This Qurʾān was decorated with silver, possessed ornamental bands in black ink serving as *sūra* dividers, and separated adjacent verses with a dot. According to the grandson, this Qurʾān was copied when ʿUṯmān had the codices written. Attempting to explain the composite nature of the codex, Cook considers the possibility that it was damaged at some point and its last fifth was recopied from a codex that had reached Medina from somewhere else. This scenario is more complex than it needs to be. Mālik b. Abī ʿĀmir was a scribe in ʿUṯmān's project of standardization. He thus had access to the *muṣḥafs* before they were sent out to the cities. If he made his personal copy before the *muṣḥafs* were dispatched, he would have had to go where these *muṣḥafs* were kept. There is no reason for expecting that in each sitting he picked up the same *muṣḥaf* to copy. See Michael Cook, "A Koranic Codex Inherited by Mālik from his Grandfather", in *Proceedings of the Sixth International Congress on Graeco-Oriental and African Studies, Graeco-Arabica*, ed. Vassilios Christides and Theodore Papadopoullos, VII-VIII, Nicosia, 7-8 (1999-2000), p. 93-105; al-Aʿẓamī, *The History of the Qurʾānic Text*, p. 97-8, 170-2.

[44] Rabb, "Non-Canonical Readings", p. 84-127; Yasin Dutton, "An Early *Muṣḥaf* According to the Reading of Ibn ʿĀmir", *Journal of Qurʾānic Studies*, 3/1 (2001), p. 71-89; *id.*, "Some Notes on the British Library's 'Oldest Qurʾān Manuscript' (Or. 2165)", *Journal of Qurʾānic Studies*, 6 (2004), p. 43-71; *id.*, "An Umayyad Fragment of the Qurʾān and its Dating", *Journal of Qurʾānic Studies*, 9/2 (2007), p. 57-87.

[45] See Dutton, "Some Notes"; Rabb, "Non-Canonical Readings". Both are cited in the last footnote. Rabb introduces some corrections to Dutton's work.

Qur'ān matches that of Kūfa.[46] Muḥammad Muṣṭafā l-Aʿẓamī has conducted similar comparisons in the latest edition of his book on the Qur'ān.[47] However, not every Qur'ān matches the data of *qirā'āt*. Naturally, the manuscripts can accumulate noise (*i.e.* changes due to a copyist's discretion, *e.g.* his choice of spelling conventions, or error) and contamination (influence from another branch of the tradition); and these could accumulate in the course of hundreds of copyings. Verse divisions are particularly susceptible to such changes. For example, a single Ḥiǧāzī folio whose image was published recently displays numerous deviations in the way the verses were divided when compared with the data in the *qirā'āt* literature.[48] Likewise, Déroche has noted that MS Arabe 328 in the Bibliothèque nationale de France (BNF) does not match neatly any of the *qirā'āt*-reported patterns.[49]

To help locate the upper text of Ṣanʿā' 1, the present essay follows in the footsteps of Dutton. I find a fit with information in the *qirā'āt* literature. It will emerge that the likelihood that this match is due to chance is nil.

The Date of the Upper Writing of Ṣanʿā' 1

Why was the lower script erased and written over? Perhaps the original owner deemed the lower writing obsolete after ʿUṯmān's standardization. This would be likely if the codex belonged to a mosque. Perhaps after the original owner died, the new owner wished to replace the text with the ʿUṯmānic version. Alternatively, it is possible that the original codex needed to be written over because it was damaged or worn out. This would fit the fact that a narrow horizontal strip along the bottom edge of some of the leaves was cut and discarded before the text was written over. The question remains open.

When was the upper layer written? Art-historical and paleographic considerations assign the upper writing to the earliest stratum of extant Qur'ān manuscripts; but, at present paleographic knowledge is not refined enough to allow dating the manuscripts in this earliest group relative to one another. Such methods establish only that the upper writing could be from the first half of the seventh century AD, the second half of that century, or (though

[46] Tayyar Altıkulaç and Ekmeleddin İhsanoğlu, *Hz. Osman'a izâfe edilen mushaf-ı şerif: Topkapı Sarayı Müzesi Nüshası*, Istanbul, İslam Târih, Sanat ve Kültür Araştırma Merkezi, 1428/2007. I owe this reference to Shahzad Bashir.

[47] Al-Aʿẓamī, *The History of the Qur'ānic Text*, p. 182-8.

[48] David Roxburgh, *Writing the Word of God: Calligraphy and the Qur'an*, The Museum of Fine Arts, Houston, 2007, p. 6-7.

[49] Déroche, *La transmission écrite du Coran*, p. 77-102. Déroche notes that the verse division patterns of the BNF manuscript are slightly closer to the Ḥimṣī tradition reported in the sources.

perhaps less likely) from early eighth century.[50] These methods thus leave any date between AD 632 and AD 700 feasible. Some might conjecture that the upper layer was written shortly after ʿUṯmān's standardization around AD 650. The fact that all the other known early manuscripts are ʿUṯmānic suggests that the standard Qur'ān became predominant quickly. So, there may have been a tendency to replace non-ʿUṯmānic Qur'āns. That said, literary evidence makes it clear that not everybody gave up non-ʿUṯmānic codices, that non-ʿUṯmānic variants were handed down by prominent scholars, and that physical copies of such codices were extant in the early centuries.[51]

Finally, one may ask *where* the upper layer was written. To help answer that question, I set out to locate the upper writing within the "reading" traditions of the cities as described in the *qirāʾāt* literature. To do so, one may use two resources: the orthographic variants of the cities and their varying systems for dividing the passages into verses.

The Provenance of an ʿUṯmānic Qurʾān: Results and Methods

Recent attempts by Yasin Dutton and others to bring the data of the *qirāʾāt* to bear on several of the earliest available ʿUṯmānic manuscripts have proved a great success. In this essay, I follow in their methodological footsteps to shed light on the provenance of the ʿUṯmānic upper writing of Ṣanʿāʾ 1.[52] The method is to compare the data of the manuscript with the limited number of patterns reported in the *qirāʾāt* literature for different cities and see if there is a match. (The *qirāʾāt* genre was defined above, on p. 25.) It emerges that the upper text could not have been written in Damascus, Ḥimṣ, Kūfa, or Baṣra, but is consistent with the traditions of Mecca or Medina. It was probably written in one of those cities, in Yemen, or in Egypt.

---

[50] On paleography and art history, see these references and the sources cited therein: Blair, *Islamic Calligraphy*, p. 101-40; Déroche, "New Evidence about Umayyad Bookhands", p. 611-42; Ṣalāḥ al-Dīn al-Munağğid, *Dirāsāt fī taʾrīḫ al-ḫaṭṭ al-ʿarabī munḏu bidāyatihi ilā nihāyat al-ʿaṣr al-umawī*, Beirut, Dār al-kitāb al-ğadīd, 1972.

[51] See *e.g.* Ibn Abī Dāwūd, *al-Maṣāḥif*, p. 14-8; Ibn al-Nadīm, Abū l-Farağ Muḥammad b. Abī Yaʿqūb Isḥāq, *Kitāb al-Fihrist*, ed. Riḍā Tağaddud, n.d., p. 29; Rāmyār, *Tārīḫ-i Qurʾān*, p. 362-4, 382-4, 389-92.

[52] The contributions of Yasin Dutton and Intisar Rabb are cited in footnote 44 above (Rabb refines the work of Dutton). For related approaches by Tayyar Altıkulaç, Muḥammad Muṣṭafā l-Aʿẓamī, and now François Déroche see footnotes 46, 47 and 49. For the broad significance of such studies, see the discussion above beginning on p. 27. To my knowledge, the earliest scholars known to have used this method are those from the second AH/eighth century AD who applied it to a manuscript belonging to the grandfather of Mālik b. Anas, for which see the secondary sources cited in footnote 43.

To appreciate just how formidable this approach can be, let us consider a simplified hypothetical example. Suppose one is interested only in differences in the way passages are divided up into verses, and suppose further that in a partial Qur'ān manuscript, there are twenty points in the text at which the *qirā'āt* literature reports disagreements over the existence of a verse separator. The total number of possible ways in which a scribe may assign verse separators to these twenty points is about a million, specifically 1,048,576 (*i.e.* 2^{20}). Yet the *qirā'āt* literature offers only seven patterns, two associated with Medina, and five with Mecca, Kūfa, Baṣra, Damascus, and Ḥimṣ. That means that the *a priori* probability that any manuscript fits one of these seven patterns at those twenty spots is virtually nil (about 0.000007). Therefore, if a manuscript does fit one of these profiles, the match is so striking that it cannot be explained by chance.

It should be noted that the *qirā'āt* were not intended to describe everything found in manuscripts, including errors; rather, they aimed to convey the knowledge of experts about what the correct text is likely to be, knowledge that was only partly based on analyses of contemporaneous manuscripts. This makes their affinity with some of the earliest available manuscripts noteworthy. One would not necessarily have expected a close correspondence. After all, in the earliest phase, many a copyist's skill set need not have exceeded literacy and the possession of good handwriting, whereas the bar for being a scholar would have been set higher.[53] Even the most meticulous scribe's text could be only as good as the manuscript he was copying, which would have incorporated the changes accumulated in the course of the hundreds of earlier copyings of careful and careless scribes.

With regard to verse divisions, in particular, a significant amount of accumulated noise (copyists' errors and exercises of discretion) eventually led to a very complex reality of which the *qirā'āt* capture but a cross-section. So, one cannot take it for granted that there will be a match—and indeed some manuscripts do *not* exhibit a close match when it comes to patterns of verse divi-

[53] There is a substantial amount of evidence for mistakes of individual copyists in Qur'ānic manuscripts (nonsense-generating haplography, dittography, verse-separator placement, spelling mistakes, etc., and singular readings). This, of course, does not mean that every scribe was careless. There is a parallelism with the New Testament textual tradition: there were plenty of sloppy copyists in the early times, although "the case of P^{75} shows that at least some scribes were capable of care" (J.R. Royse, "Scribal Tendencies in the Transmission of the Text", in *The Text of the New Testament in Contemporary Research: Essays on the Status Quaestionis*, ed. B.D. Ehrman and M.W. Holmes, Grand Rapids, 1995, p. 239-52). Over time standards improved; thus, "early scribes did not exercise the care evidenced in later transcriptions" (*ibid.*). For two instances of scribal incompetence so egregious as to be entertaining, see Bruce M. Metzger, *The Text of the New Testament: Its Transmission, Corruption, and Restoration*, 3rd edition, Oxford, 1992, p. 194-5.

sion. And yet, remarkably, there are instances of a close correspondence between manuscripts and the *qirāʾāt*. In such cases, the *qirāʾāt* can help reveal the provenance of manuscripts.

How to Compare Manuscripts and the Qirāʾāt ("Readings") Literature

What kinds of variants can be found in manuscripts, and which ones are treated in the *qirāʾāt* genre? ʿUṯmānic Qurʾān manuscripts could encompass several different types of variants: differences in vowel markings and *hamza* markings in those manuscripts that do have such markings, variations in the small dashes distinguishing one consonant from another one of a similar skeletal shape (such as the letter ب versus ٮ) (cf. first row of Table 5), spelling differences that do not change a word or its pronunciation, spelling differences that *do* change the pronunciation of a word but not its meaning (*e.g.* اولا الالبب vs. اولى الالبب), and the use of different words that have similar shapes. Among these types, some (such as those involving diacritics and vowel markings) can be recognized less frequently or not at all in the earliest manuscripts. For example, the upper writing of Ṣanʿāʾ 1 does not have vowel markings.

The *qirāʾāt* literature provides systematic information on all the above types of variants, but it is not as thorough for graphical (or skeletal, *rasm*) variants that neither potentially change one word into another nor make a difference in its pronunciation (second row of Table 5). In the English language, such variations would be akin to the difference between the spellings "adviser" and "advisor", or between "behavior" and behaviour". Though understandably relatively marginal in the *qirāʾāt* literature, variants of this type constitute the largest category of variations in the manuscript evidence. These variations owe, most of all, to the lack of consistency in the way the long vowels were represented in early Arabic writing, as different copyists used different spelling practices, and sometimes the same scribe switched from one spelling of a word to another.[54] Among such orthographic variants, by far most involve the long vowel *ā*. For example, the *ā* sound in the word *ʿalā* could be represented by the letter *alif* (ا) or the letter *yāʾ* (ى), yielding either علا or على as alternative spellings. More often, however, the choice was between denoting the *ā* sound by the letter *alif* and simply not representing it. Thus the morpheme *kātib* (in Kor 2, 283) could be written as *k - ā - t - b* (كاتب) or *k - t - b* (كتب).

[54] *E.g.* al-Aʿẓamī, *The History of the Qurʾānic Text*, p. 145; for examples from early non-Qurʾānic manuscripts, see *Ibid.*, p. 163-4.

Table 5. ***The utility of the literary sources for determining the provenance of a manuscript fragment in the ʿUṯmānic textual tradition depends on the types of variants found in the fragment.***

Type of variant	Presence in Manuscripts	Information from the Literary Sources	Usefulness of *Qirāʾāt* for Assigning Provenance
1. Different vowels or different consonant-distinguishing diacritics	Earliest manuscripts have no or a limited number of diacritics and vowel marks	Extensive	Not useful for manuscripts that lack diacritics or vowel marks at spots of interest
2. Skeletal differences that do not necessarily change pronunciation	Very frequent	Limited	Not very useful
3. Skeletal differences that necessarily change pronunciation ("skeletal-morphemic")	Frequent to a degree consistent with written transmission	Precise & accurate for the original codices sent out by ʿUṯmān (Mecca, Medina, Kūfa, Baṣra, Syria). Lacking vis-à-vis changes accumulated later.	Useful if some of the 36 spots where the original ʿUṯmānic codices disagree are included in the manuscript fragment
4. Differences in the locations of verse divisions	Frequent	Lacking for the original codices sent out by ʿUṯmān. *Some* of the later branches of the tradition are recorded (Mecca, Medina, Damascus, Ḥimṣ, Kūfa, Baṣra). Not all later schemes are recorded.	Useful if the verse divisions in the manuscript match one of the systems recorded for the systems. (There is not always a match.)

Note that in the second case (*k-t-b*), with the *alif* absent, without the linguistic context the reader would not know whether to read the term as *kataba*, *kutiba*, *kutub*, *kitāb*, *kātib*, or *kuttāb*, all seemingly valid choices; but here (Kor 2, 283) the context removes ambiguity. In the case of the first spelling (*k-ā-t-b*), by contrast, the orthography itself fixes the first vowel as *ā*. There are other cases where even the context does not completely disambiguate a word. In these instances, the *qirā'āt* literature gives the variations in pronunciation that may have existed among the early authorities.

What types of variation can help one check manuscripts against the *qirā'āt* "readings"? The type of variants about which the *qirā'āt* have the most to say, namely those concerning different vowellings of the skeletal text, is of the least use for unvowelled manuscripts. Likewise, the largest category of variants found in manuscripts, namely spelling variations that do not necessarily change a word (*e.g.* *k-ā-t-b* vs. *k-t-b* for *kātib* in Kor 2, 282), is not very useful, because the *qirā'āt* literature does not give a thorough description of regional spelling practices. In any case, it is highly likely, given that spelling conventions were not standardized, that different spellings of a word did not *usually* signify regional variation. For example, two scribes from the same time and place may spell the word *'alā* differently (على vs. علا), as did the two copyists of the upper layer of Ṣan'ā' 1. (Scholars did, however, record over 200 instances in which codices of the main cities agreed on the spelling of a word.[55] They also recorded some cases in which the codices of one city seemed to agree.[56] This suggests that further research might identify certain words of which the spelling could have diagnostic value.)

This leaves mainly two types of variants: (1) skeletal-morphemic differences, (2) differences in the way verses are demarcated. I take these up in turn.

Skeletal-Morphemic Variants in the 'Uṯmānic Textual Tradition

How should one go about comparing the skeletal text (*rasm*) of a manuscript with the *qirā'āt*, and what can one learn from such a comparison? This section considers the relatively small subset of such variants that *necessarily* change a word, even if only slightly, or make a difference in the way it is pronounced. It is convenient to have a label for them; and so I call them "skeletal-morphemic variants". In other words, the present discussion is limited to the skeletal written form, and excludes orthographic variations that do not

[55] For Ibn Abī Dāwūd al-Siğistānī's list of words of which the spelling was agreed upon, see his *Kitāb al-Maṣāḥif*, p. 105-17. As could be expected, relatively few of these involve the letter *alif*. This confirms that the use of *alif* was inconsistent.

[56] *E.g.*, for Medina, *Ibid.*, p. 104.

necessarily change a word or its pronunciation (third row of Table 5). The use of *fa-* instead of *wa-* is an example of a skeletal-morphemic variation; but differences in the spelling of a word—*e.g.* the spelling *k - t - b* (كتب) instead of *k - ā - t - b* (كاتب) for *kātib* in Kor 2, 283—are not.

As discussed above (p. 26), Cook has shown that the *qirāʾāt* literature provides authentic and accurate information about differences of this kind among the master codices that ʿUṯmān sent to the cities.[57] If the ʿUṯmānic textual tradition is one that evolved over time, we have the splendid fortune of being able to say that we already know the initial state in this chain of succession; we know the roots and the trunk in the genetic tree diagram, the original ʿUṯmānic master codices. This holds, though, only at the skeletal-morphemic level, not always for the less important cases where an orthographical difference does not necessarily make a difference in pronunciation. For example, this often does not hold for the spellings of words. In many cases, we do not know the original spellings of words in the codices ʿUṯmān sent out. Most manuscripts, on the other hand, constitute evidence of the later developments of the written text: namely, the leaves and branches of the ʿUṯmānic textual tradition. For once, then, scholars need not toil to reconstruct a prototype from the later realizations. Thanks to the efforts of the earliest Qurʾān specialists, one is not left at the mercy of the copyists and the changes they introduced in the texts.

These insights should make it clear how to approach an ʿUṯmānic manuscript. If its orthography matches one of the profiles reported for Mecca, Medina, Kūfa, Baṣra, or Syria, then the manuscript belongs to that city or to regions following the tradition of that city. On the other hand, words inconsistent with the regional master codices sent by ʿUṯmān can be safely regarded as later accretions. Such deviations, of course, can be handed down in a textual tradition, and therefore may have value for determining the provenance

---

[57] Michael Cook, "The Stemma of the Regional Codices of the Koran", p. 89-104. Where did the knowledge of the skeletal-morphemic differences of the regional ʿUṯmānic archetypes come from? Investigators may have gleaned the differences between regional codices by examining the original master copies of ʿUṯmān. Al-Aʿẓamī makes a case for this scenario (*The History of the Qurʾānic Text*, p. 167). The details of this early scholarly activity merit further inquiry. As another possibility, one ought to investigate whether the reported orthographical differences were derived from the *qirāʾāt*, the actual scholarly regional traditions of reciting the Qurʾān. Orthographical variants that made a difference to pronunciation would have been implicit in the *qirāʾāt* (cf. Muḥammad Sālim Muḥaysin, *al-Fatḥ al-rabbānī fī ʿalāqat al-qirāʾāt bi-l-rasm al-ʿuṯmānī*, Saudi Arabia, Ǧāmiʿat al-Imām Muḥammad b. Saʿūd al-islāmiyya, 1415/1994). As a third possibility, scholars may have conducted rudimentary textual analysis of high-quality first-century manuscripts from different cities. Whichever method was used, Cook's study shows that it yielded accurate results.

of a manuscript. This is because one may more effectively assign a manuscript to a place if one has regional profiles based on all textual characteristics, including accumulated errors. However, the current state of scholarship is not sufficiently advanced to allow one to put accumulated errors to such a use.

Verse Divisions in the ʿU̱tmānic Textual Tradition

By verse-division variants, I mean differences in the ways in which different cities or individuals divided the *sūra*s into verses. What I present here is what the literary sources (*qirāʾāt* literature) have to say about the variations in such divisions, not necessarily what one always observes in manuscripts—manuscripts can possess divisions not described in the *qirāʾāt* literature or lack verse divisions reported therein, and such deviations may be few or many. One may compare the data of a manuscript with the reported verse division schemes to see if there is a match (last row of Table 5).

The literary sources (*qirāʾāt* reports) describe some of the traditions of dividing verses in Mecca, Medina (two traditions), Kūfa, Baṣra, Damascus, and Ḥimṣ. These traditions may agree or disagree at a particular point. In thirty-eight of the *sūras*, no such difference arose: the cities were unanimous on how to divide the *sūras* into verses. Ten *sūra*s contain a single disputed division apiece, while most of the rest of the *sūra*s have several disputed divisions. *Sūra* 20 stands out with over twenty disputed cases. Interestingly, the density of disputed points is greater in *sūra*s with shorter verses. The similarities and differences among the verse-division schemes reported in the *qirāʾāt* reveal something about how they came about:

On the one hand, the reported verse-division variants are fewer than what one would expect if different cities had divided the *sūra*s independently. Despite the fact that some verse-endings are "obvious" cases, it seems that if two persons independently used discretion to insert verse divisions in the undivided text, the resulting disparity would be great compared to the differences among the schemes reported in the *qirāʾāt* literature. This shows that the verse-division traditions are linked genetically: they were derived from one another and/or from common sources. Furthermore, divisions that are not disputed, *i.e.* the vast majority of divisions, can be safely ascribed to a prototype.

On the other hand, verse-division variants are much more numerous than the skeletal-morphemic differences of the regional ʿU̱tmānic archetypes. The gap might reflect, in part, the fact that the reports on orthography describe the *original* codices that ʿU̱tmān disseminated, whereas the reports on verse division concern the regional traditions that had evolved from those

codices. There is, in other words, a time gap allowing, in principle, for gradual divergence from the original state. However, this alone does not explain the large number of variants. The more important factor is that verse divisions inherently were more susceptible to individual discretion, transmission error, and/or cross-contamination. The scribes must have considered the skeletal-morphemic text to be more important than the way in which it was divided up, leading them to exercise less care in transmitting the separator marks exactly as they found them.

The presence of accumulated "noise" (transmission error and exercise of individual discretion) and "contamination" (the influence of one city's textual tradition upon another) in the verse-division systems described by the *qirāʾāt* makes it relatively difficult to reconstruct the genetic relationships between the verse-division practices of the cities. (This is in contrast to the straightforward manner in which such relationships can be detected based on orthographic variants.)[58] While such a reconstruction falls outside the scope of the present essay, two unmistakable patterns may be pointed out: First, Mecca, Medina I, and Medina II fall into a cluster apart from Kūfa, Baṣra, and Syria. (Syria includes Damascus and Ḥimṣ, whose systems were relatively close to each other; and Medina I and II refer to two different systems in Medina.) The verse division systems of any two cities in this cluster are closer to each other than to that of any city not in the cluster. Second, surprisingly, Medina I is closer to Mecca than to Medina II. Similarly, Medina II is closer to Mecca than to Medina I.[59]

The greater number of verse division variants (as compared to skeletal-morphemic variants) makes them a particularly useful diagnostic tool in cases in which only a small part of a codex has survived. This is because, on average, a passage contains more disputed verse divisions than disputed features of the skeletal text. Such variants can be very helpful in pinning down the provenance of a manuscript. For example, the fact that the British Library's "oldest Qurʾān manuscript" has been assigned to the tradition of Ḥimṣ shows that a high level of discrimination is possible, given the closeness between the verse demarcation schemes of Ḥimṣ and Damascus.[60]

[58] *Ibid.*

[59] There are four possible stemmata: (1) M2 ← Mec ← M1; (2) M1 ← Mec ← M2; (3) (M1 ← Mec) and (M2 ←Mec); (4) (Mec ← M1) and (Mec ← M2). This last is unlikely, since it does not explain why M1 and M2 belong to the same cluster. Here, M1, M2, and Mec refer respectively to Medina I, Medina II, and Mecca.

[60] See the sources cited in footnote 45.

Verse Divisions of the Upper Text of Ṣanʿāʾ 1

By checking for the presence or absence of disputed verse separators, I now determine if the upper text matches any of the verse division traditions reported in the literary sources.

Sūra 2, 265-277[61] and 2, 277-286[62]

These passages do not help pin down the text. None of the eleven disputed verse endings reported by al-Dānī or others for *sūra* 2 occur in the passages found in the leaves.[63]

The upper text has two verse divisions not reported in al-Dānī, at *wa-lā tayammamū l-ḫabīṯa minhu tunfiqūna* (2, 267) and at *wa-l-muʾminūna* (2, 285). The ten-verse marker designating verse 270 appears immediately after *wa-llāhu bi-mā taʿmalūna ḫabīrun* (2, 271). There is another ten-verse marker at *lā yuẓlamūna* (2, 281). One could try to see how the leaf matches up with the verse totals reported for the cities, but given the unusual length of the *sūra* and the two anomalous additional verse divisions already observed, chances are good that the manuscript would have possessed (or lacked) another verse division anomalously, making any match misleading.

---

61 Stanford '07.

62 The David folio.

63 The disputed points in *sūra* 2 are as follows:

alif lām mīm (2, 1): Kūfa alone counted it as a verse ending;
ʿaḏābun alīmun (2,10): Syria alone counted it;
innamā naḥnu muṣliḥūna (2, 11): Syria alone did *not* count it as a verse ending;
illā ḫāʾifīna (2, 114): Baṣra alone counted it;
yā ūlī l-albābi (2, 197): Medina I and Mecca alone did *not* count it;
mā lahu fī l-āḫirati min ḫalāqin (2, 200): Medina II alone did *not* count it;
māḏā yunfiqūna (2, 219): Medina I and Mecca alone counted it;
laʿallakum tatafakkarūna (2, 119): Medina II, Kūfa, and Syria alone counted it;
qawlan maʿrūfan (2, 235): Baṣra alone counted it;
al-ḥayyu l-qayyūmu (2, 255): Medina II, Baṣra, and Mecca alone counted it;
min al-ẓulumāti ilā l-nūri (2, 257): Medina I alone counted it.

The *sūra* has 285 verses according to Medina, Mecca, and Syria, 286 for Kūfa, and 287 for Baṣra. See al-Dānī, *al-Bayān fī ʿadd āy al-Qurʾān*, Kuwait, Dār al-našr, 1414/1994, p. 88, 89, 91, 95-6, 100-1, 108, 140-2. Cf. Anton Spitaler, *Die Verszählung des Koran nach islamischer Überlieferung*, Munich, Verlag der Bayerischen Akademie der Wissenschaften, 1935, p. 32-4.

Sūra 6, 49-61[64] and 6, 149-59[65]

None of the four disputed verse endings reported by al-Dānī for *sūra* 6 occur in the passages found in the leaves.[66] Fortunately, the verse count implied by the ten-verse markers in the upper text sheds light on the verse divisions. There are ten-verse separators at *yafsuqūna* (6, 49), *mubīnin* (6, 59), and *aǧmaʿīna* (6, 149), indicating fifty, sixty, and 150 verses respectively. These numbers would result if there were a division at *al-ẓulumāti wa-l-nūra* (6, 1). (This is the only point disputed in this part of the *sūra*, the next one coming at 6, 66.) Accepting the presence of a division at *al-ẓulumāti wa-l-nūra* (6, 1), one may ask how this division scheme lines up with the city profiles. The division at *al-ẓulumāti wa-l-nūra* fits a Meccan or Medinan provenance and rules out Kūfa, Baṣra, and Syria.[67] This conclusion rests on the assumption that there are no anomalous verse divisions in the manuscript prior to 6, 49. This is not an unreasonable assumption: the fact that the three ten-verse markers all appear at the expected positions suggests regularity.

Sūra 20, 1-43[68]

It is fortunate that one image covers the first forty-three verses of *sūra* 20, since this *sūra* has the largest number of disputed verse endings in the entire Qurʾān—over twenty of them.[69] Of these the folio image covers nine. Damas-

[64] Image 152255B, UNESCO CD.

[65] Image 152254B, UNESCO CD.

[66] The disputed points in *sūra* 6 are as follows:

wa-ǧaʿala l-ẓulumāti wa-l-nūra (6, 1): Medina I & II and Mecca alone counted it as a verse ending;
qul lastu ʿalaykum bi-wakīlin (6, 66): Kūfa alone counted it;
kun fa-yakūnu (6, 73): Kūfa alone did *not* count it as a verse ending;
ilā ṣirāṭin mustaqīmin (6, 161): Kūfa alone did *not* count it as a verse ending;
huwa llaḏī ḫalaqakum min ṭīnin (6, 2): a report says that the Medinans (*ahl al-Madīna*), unlike the rest, counted this as a verse ending.

The total verse count is 165 for Kūfa, 166 for Baṣra and Syria, and 167 for Medina and Mecca. See al-Dānī, *al-Bayān*, p. 89, 91-2, 99, 151-4. Cf. Spitaler, *Verszählung*, p. 36-7.

[67] There is a verse separator at *yaṣdifūna* in 6, 157, but it looks like a scribal error. It appears awkwardly mid-sentence. It was apparently confused with the real verse ending, which is at another occurrence of the word *yaṣdifūna*. The scribe's eye appears to have leaped to a different occurrence of the same word, a common type of scribal error.

[68] Image 152256B in UNESCO CD.

[69] The disputed verse endings for *sūra* 20 are as follows:

ṭāhā (20, 1): Kūfa alone counted it as a verse ending;
nusabbiḥaka kaṯīran (20, 33) and *wa-naḏkuraka kaṯīran* (20, 34): Baṣra alone did *not* count these two spots as verse endings;
fa-qḏifīhi fī l-yammi (20, 39): Ḥimṣ alone counted it;

cus is ruled out four times—by the absence of verse separators at *taqarra ʿaynuhā wa-lā taḥzan* (20, 40), *wa-fatannaka futūnan* (20, 40), *fī ahli Madyana* (20, 40), and *wa-ṣṭanaʿtuka li-nafsī* (20, 41). Ḥimṣ is ruled out four times—by the absence of verse separators at *fa-qḏifīhi fī l-yammi* (20, 39), *wa-fatannaka futūnan* (20, 40), and *wa-ṣṭanaʿtuka li-nafsī* (20, 41), and by the presence of a verse separator at *maḥabbatan minnī* (20, 39). Kūfa is ruled out three times—by the absence of a verse separator at *Ṭāhā* (20, 1) (which can be inferred by the location of the ten-verse marker)[70] and *wa-ṣṭanaʿtuka li-nafsī* (20, 41), and by the presence of a verse separator at *maḥabbatan minnī* (20, 39). Baṣra is ruled out three times—by the presence of verse separators at *naḏkuraka kaṯīran* (20, 34) and *maḥabbatan minnī* (20, 39), and by the absence of a verse separator at *wa-fatannaka futūnan* (20, 40).[71] [72]

---

maḥabbatan minnī (20, 39): Kūfa, Baṣra and Ḥimṣ alone did *not* count it as a verse ending;
kay taqarra ʿaynuhā wa-lā taḥzan (20, 40): Damascus (and in one narration, Ḥimṣ) alone counted it;
wa-fatannāka futūnan (20, 40): Baṣra, Damascus, and Ḥimṣ alone counted it;
fī ahli Madyana (20, 40): Damascus (and in one narration, Ḥimṣ) alone counted it;
wa-ṣṭanaʿtuka li-nafsī (20, 41): Kūfa, Damascus, and Ḥimṣ alone counted it;
maʿanā banī Isrāʾīla (20, 47): Damascus and Ḥimṣ alone counted it;
wa-laqad awḥaynā ilā Mūsā (20, 77): Damascus (and in one narration, Ḥimṣ) alone counted it;
min al-yammi mā ġašiyahum (20, 78): Kūfa alone counted it;
waʿdan ḥasanan (20, 86): Medina II (and in one narration, Damascus) alone counted it;
ġaḍbāna asifan (20, 86): Medina I and Mecca (and in one narration, Ḥimṣ) alone counted it;
fa-kaḏālika alqā l-sāmirī (20, 87): the second Medinan (and in some narrations, Damascus and Ḥimṣ) alone did *not* count it;
wa-ilāhu Mūsā (20, 88): Medina I and Mecca alone counted it;
fa-nasiya (20, 88): Medina I and Mecca alone did *not* count it;
ilayhim qawlan (20, 89): Medina II (and in some narrations, Damascus and Ḥimṣ) alone counted it;
iḏ raʾaytahum ḍallū (20, 92): Kūfa alone counted it;
qāʿan ṣafṣafan (20, 106): Medina I & II and Mecca alone did *not* count it;
minnī hudan (20, 123): Kūfa (and in some narrations, Ḥimṣ) alone did *not* count it;
maʿīšatan ḍankan (20, 124): Ḥimṣ alone counted it;
zahrata l-ḥayāti l-dunyā (20, 131): Kūfa (and in one narration, Ḥimṣ) alone did *not* count it.

The *sūra* has 132 verses according to Baṣra, 134 according to Mecca and Medina I & II, 135 according to Kūfa, 138 according to Ḥimṣ, and 140 according to Damascus. See al-Dānī, *al-Bayān*, p. 88-108, 183-6; Spitaler, *Verszählung*, p. 44-7.

70 There are ten-verse separators at *yā Mūsā* (20, 11), *al-ūlā* (20, 21), *azrī* (20, 31), and *yā Mūsā* (20, 40).

71 Due to the poor quality of the image, I cannot determine whether there is a verse separator at *nusabbiḥaka kaṯīran* (20, 33). One would expect a verse ending there to match the one at *wa-naḏkuraka kaṯīran* (20, 34).

72 Where narrations about a city disagree, I disregard the city, yielding neither a match nor a mismatch.

The presence of ten-verse markers at *yā-Mūsā* (20, 11), *al-ūlā* (20, 21), and *azrī* (20, 31) to mark the tenth, twentieth, and thirtieth verses is compatible with all the cities except Kūfa. However, the presence of a ten-verse separator at *yā-Mūsā* (20, 40) to mark verse forty is compatible only with Kūfa, Mecca, and Medina.

Sūra 31, 24-34[73]
There is a ten-verse marker at *šakūr* (31, 31), indicating the thirtieth verse. This is compatible only with Mecca and Medina, ruling out Kūfa, Baṣra, and Syria. There is no verse separator at *lahu l-dīna* (31, 32). This is compatible with Kūfa, Medina, and Mecca, ruling out Baṣra and Syria. In sum, Baṣra and Syria are ruled out twice, and Kūfa once. The verse divisions fit Mecca and Medina.[74]

Sūra 32, 1-4[75]
There is no separator at *alif-lām-mīm*. Kūfa alone puts a separator there, and is thus excluded.

Sūra 59, 4-10[76]
There are no disputed verse endings in *sūra* 59, and the small leaf fragment does not offer any variants.[77]

Synopsis

None of the disputed points distinguishing between Mecca, Medina I, and Medina II happens to fall in the passages available for this study. But these passages include plenty of disputed points that could decide between these three Ḥiǧāzī traditions taken together, on the one hand, and the cities outside the Ḥiǧāz, on the other hand. As one examines these instances, a striking pattern emerges: *There are fourteen cases in which the presence or absence of a verse separator could have ruled out the Ḥiǧāz in favor of some other locale(s), yet in every single such case the Ḥiǧāz passes the test.* The probability of this pattern occurring by chance is near zero. Ruled out conclusively are Kūfa, Baṣra, Damascus, and Ḥimṣ. These cities fail the test every single time the alternative is between them and the Ḥiǧaz (six times for Kūfa, seven times for Baṣra,

[73] *Maṣāḥif Ṣanʿāʾ*, no. 4, p. 59 (English) = 44 (Arabic).
[74] Al-Dānī, *al-Bayān*, p. 91, 106, and 206; Spitaler, *Verszählung*, p. 51.
[75] *Maṣāḥif Ṣanʿāʾ*, no. 4, p. 59 (English) = 44 (Arabic).
[76] Image 152257-B in UNESCO CD.
[77] See al-Dānī, *al-Bayān*, p. 243; Spitaler, *Verszählung*, p. 30.

and eight times for Syria). There is thus a stunning match between the *qirā'āt* literature and manuscript evidence.

III. The Lower Writing of Ṣanʿāʾ 1

Statement of the Problem

The lower writing is about as old as the parchment on which it appears (see above, p. 12). And the words and phrases it conveys are *at least* as old as the writing. The parchment, according to radiocarbon dating, has a 91.8 % chance of dating from before ʿUṯmān's death in AD 656, and a 95.5 % chance of dating from before AD 661. It is almost certainly older than AD 671 (probability 98.8 %). Most likely, it was produced no more than 14 years after the death of the Prophet Muḥammad: the probability of this is 75 %, or a three-to-one likelihood. It is even slightly more likely than not that it was produced within four years after the Prophet's death (probability 56.2 %). (For more details, see Table 1, above, p. 11.) What makes it even more likely that Ṣanʿāʾ 1 was created in the first half of the seventh century is that after ʿUṯmān's standardization *ca.* AD 650, copying non-standard Qur'āns became less common. Those who defied the edict and held on to their Companion codices were a minority. The standard version quickly became predominant, as shown by the manuscript record and literary sources.

The lower writing, in any case, dates from the era of the Companions of the Prophet. It was thus an early copy of a Companion C-1's codex. So far, this would make the C-1 textual tradition contemporaneous with the ʿUṯmānic tradition, in the sense that they both date, at the latest, from the era of the Companions. There is more to be said, however. In this essay I seek to determine whether the wording of one codex can be determined to be even older—in the sense of being either the source of the other or a significantly more accurate copy of the common ancestor of both.

Two approaches to this question may be readily dismissed as callow. First, it would be simpleminded to say, because C-1 may be labeled "pre-ʿUṯmānic", that its wording necessarily is older than that of the ʿUṯmānic text type. If by "pre-ʿUṯmānic" one means "predating ʿUṯmān's act of standardization", then the ʿUṯmānic text type may be "pre-ʿUṯmānic" too. The wording of the codices ʿUṯmān sent out may have predated the standardization, just as ʿUṯmān, the people he set to the task, and the material they worked with existed before AD 650. To ward off any prejudging of the matter, I call C-1 and other Companion textual traditions "non-ʿUṯmānic" rather than "pre-ʿUṯmānic".

Second, it is equally naive to say, because the upper writing came after the lower one in this particular manuscript, that the wording of the ʿUṯmānic Qurʾān must be later than that of C-1. That is tantamount to assuming that when the upper writing appeared, it represented the totality of the ʿUṯmānic tradition, meaning that the wording of ʿUṯmān's codex did not exist before it appeared on this particular manuscript. But the upper writing must have been one of innumerable ʿUṯmānic manuscripts in circulation, a fact supported by the manuscript record, literary evidence, and common sense.[78] It is also unlikely that the lower writing represented the beginning and the end of the C-1 textual tradition. In other words, one must distinguish between a book as a physical object and a book as a text, for the wording is usually older than the manuscript, and text types are almost always older than texts. That one codex was produced after the other does not mean that its wording came later. The upper layer of Ṣanʿāʾ 1, rather than capturing the *genesis* of the ʿUṯmānic textual tradition, may represent its *encroachment* upon a parallel, contemporaneous C-1 textual tradition. At least, that possibility cannot be ruled out before analysis.

The fundamental question, therefore, remains wide open: what is the relationship between the *text types* of ʿUṯmān and C-1? Is it possible to determine if one is older than the other? This question, of course, calls for textual criticism and a model of textual transmission. I will argue that the model that best explains the origin of the differences between the textual traditions of C-1 and ʿUṯmān is dictation, though not word-by-word dictation. Imagine a Companion of the Prophet doing his best to write down a Qurʾān ("Recitation") that was being recited faster than he could record with perfect accuracy. The Companion succeeded in getting most of the phrases right, thus explaining the commonalities of the various Companion codices and ʿUṯmān's text type; yet, naturally, some differences arose. In due time, the Companion C-1's text was copied, and perhaps the copies were copied, creating a C-1 textual tradition and a family of closely related texts constituting the C-1 text type, of which the lower writing of Ṣanʿāʾ 1 is one instance. In this scenario, the repeated copyings did not necessarily involve dictation, but the C-1 archetype itself resulted from dictation.

The textual differences between C-1 and the standard Qurʾān are strikingly similar to the variations that tended to arise in *ḥadīṯ* and *āṯār* texts in the late second and early third century AH. By then, the use of writing was universal, and standards of fidelity in transmission were approaching their peak. Yet occasional differences arose between two versions of a text dictated by a com-

[78] The extant manuscript fragments from the first century, which tend to be large, sturdy, and expensive, must be but a fraction of early high-quality manuscripts; and expensive, high-quality manuscripts must have been a small fraction of all manuscripts.

mon source. Such variations included not only those familiar from copyists' errors—many such errors can occur during dictation, too—but also differences more typical of the oral and aural aspects of transmission by dictation. Insights gathered from experience with this stage of *ḥadīṯ* transmission can be brought to bear on the study of Companion Qur'āns. Thus, for example, by comparing the variants of a *ḥadīṯ* one finds that the most common words are the most unstable, for they are the least memorable. In English, it would be as if "say" and "tell" were more interchangeable than "perdure" and "endure", or "Mecca" and "Medina". Moreover, in the *ḥadīṯ* literature, function morphemes are the most variable. The English equivalents would be words like "of", "or", "so", "by", "to", "with", "it", "and", "then", "some". The most telling variations in *ḥadīṯ*s, perhaps, involve the replacement of a word with a similar-sounding word or changes of syntax that do not change the meaning. Such changes occur less often in written transmission.[79]

The Impact of Semi-Orality: Patterns Governing the First Tier of Variations in Transmission by Dictation

One sign of dictation—or dictation in which the outcome is not subsequently checked against the original—is that the changes are more frequent than in written transmission.[80] In both forms of transmission, the scribe's memory plays a role. It holds the words between the seeing/hearing and the writing. However, memory usually plays a more profound role in dictation. This is so for several reasons: Unless the dictation takes place slowly, the scribe has to do two things at once: listen to new words and write down previous ones, making it easier to make mistakes. Moreover, in dictation, the scribe's glance does not fall on writing that can correct his memory. Furthermore, in dictation a scribe is more likely to copy sentence-by-sentence or phrase-by-phrase than word-by-word, "which makes it easier to forget the exact words".[81] The more rapidly the text is recited, the greater the strain is on the scribe's memory, and the larger and more numerous the resulting differences.

Aside from the larger number of variants, there is also a qualitative difference: each type of transmission generates certain types of changes in different proportions. Some types of variants are more common in dictation and others

---

[79] I base these conclusions on the study of numerous of *āṯār* variants, in which I compared *isnād*s with *matn*s, as part of a current book project on early Islam.

[80] For short, "dictation" in this essay generally refers to the type of dictation in which the writing is not later corrected against the manuscript that was dictated. When the text is corrected against the original, either visually or by reciting the text back, many of the changes are eliminated.

[81] The point about copying phrase-by-phrase was brought to my attention by Robert Waltz (private communication, 2009).

more common in written transmission. For example, dictation eliminates certain nonsense-generating errors that are due to the confusion of the eyes, such as haplography, dittography, and errors due to the scribe's eye leaping from a word to a similar word elsewhere on the page; but nonsense-producing errors of the scribe's hand can still happen.[82]

Changes arising from dictation tend to follow certain patterns that accord with the limitations of human memory. It is convenient to refer to the *most basic* of these changes as "first-tier". First-tier variations are the most probable, and therefore most common, changes introduced by the limitations of memory. Such changes can appear even through reasonably careful dictation. However, changes exceed the first tier if the scribe exercises less care, if the recitation takes place more rapidly, or if the time interval between the oral statement and its writing is increased.[83] In concrete terms, first-tier variations may be discussed under five headings:

1. *Changes of Minor Elements.* The *most commonly used* elements of language are what I refer to, strictly as a shorthand, as "minor". These are the easiest class of morphemes to add, drop, or substitute, for they are the least memorable. Minor elements include function morphemes like particles, articles, conjunctions, prepositions, pronouns, and some prefixes and suffixes. Minor elements also include a morpheme such as *Allāh* which occurs 2,689 times in the Qur'ān. Among the most variable elements are *wa-* and *fa-*. While such changes do appear in written transmission, in dictation their absolute numbers are greater.

Appendix 1 shows that in C-1 there are 35 minor differences with ʿUṯmān, including 12 omissions, 19 substitutions (8 involving personal pronouns),

[82] Some of the changes that are typical of dictation can be classified under what Metzger calls "errors of the mind". With regards to New Testament manuscripts, Metzger writes, "The category of errors of the mind includes those variations which seem to have arisen while the copyist was holding a clause or a sequence of letters in his (somewhat treacherous) memory between the glance at the manuscript to be copied and the writing down of what he saw there" (Metzger, *Text*, p. 192). He documents errors of the mind involving substitutions of synonyms, word order, transpositions, and assimilation of parallels (*ibid.*, p. 192-3). Note that while memory plays a role in both written transmission and dictation, it plays a bigger role in the latter, which explains why errors of the mind are more common in dictation.

[83] The *Ḥadīṯ* literature furnishes ample evidence of the full spectrum of transmission behaviors, from purely oral to purely written. Changes in *ḥadīṯ*s fit the profile of nearly complete orality in part of the first century AH, reduce to high-tier changes by the end of that century (signifying hybrid oral-written transmission), and then turn into patterns consistent with first-tier or written transmission by the end of the second century AH. These shifts occur at different rates in different circles of transmitters. It should also be noted that, in some stages, *ḥadīṯ* transmission could include hybrid oral-written forms of copying other than dictation, *e.g.* note taking.

and four additions. It should be noted that there is a close match between C-1 and ʿUṯmān, and that differences are the exception rather than the rule even when it comes to minor elements. This leaves no doubt that transmission involved writing (as in dictation) rather than being purely oral.

The remaining principles, below, apply primarily to elements that are *not* minor, thus labeled "major". In dictation, major changes form a larger proportion of variations than they do in written transmission. This is an obvious consequence of the greater role played by memory in dictation.

2. *Omissions of Major Elements.* We more frequently forget a thing or remember a thing differently than we "remember" something we have not heard at all. Thus, words are dropped (or changed) more easily than they are added. Furthermore, it is easier to forget a word if it is an item on a list. The greater likelihood of omissions as compared to additions is a feature that dictation shares with written transmission.[84] In general, if a variant is minor, or if it is an omission, then no special explanation is needed as to how it could have come about. If it is both, its occurrence is even more understandable. By contrast, if a variant is an addition, then it cannot be considered first-tier a priori, and one must ask whether there are first-tier mechanisms that might explain how it came about, mechanisms such as auto-contamination.

3. *Auto-contamination.* Auto-contamination refers to the influence, within one textual tradition, of one part of the Qurʾān on another part. This is to distinguish it from "cross-contamination", which refers to one textual tradition,

[84] In the case of dictation, at least, it is plausible that omissions occur more easily than additions. But omissions are somewhat more common even in other forms of transmission. New Testament scholars used to believe that, *ceteris paribus*, the shorter of two readings is more likely to be the original. This idea was challenged early in the twentieth century by the work of A.C. Clark, who found that omissions are more common in the manuscripts of Latin and Greek classics. This led to divided opinions in the field. See Metzger, *Text*, p. 161-3; cf. Burnett Streeter, *The Four Gospels: A Study of Origins*, 4th ed. rev., London, Macmillan, 1930, p. 131; following Clark, Streeter states that "the error to which scribes were most prone was not interpolation but accidental omission".

More recently, studies of singular readings in papyri have confirmed repeatedly that omissions were *more common*. See J.R. Royse, "Scribal Habits in the Transmission of New Testament Texts", in *The Critical Study of Sacred Texts*, ed. W.D. O'Flaherty, Berkeley, Graduate Theological Union, 1979, p. 139-61; J.R. Royse, "Scribal Tendencies", p. 239-52; Peter M. Head, "The Habits of New Testament Copyists: Singular Readings in the Early Fragmentary Papyri of John", *Biblica*, 85 (2004), p. 399-408.

Singular readings are those that are unique to a manuscript. They subsume a disproportionately large number of changes introduced by copyists and can thus be used to study the habits and idiosyncrasies of scribes. For a discussion of methodology, see, in addition to the sources cited above, E.C. Colwell, "Scribal Habits in Early Papyri: A Study in the Corruption of the Text", in *The Bible in Modern Scholarship: Papers read at the 100th Meeting of the Society of Biblical Literature*, ed. J.P. Hyatt, Nashville, Abingdon Press, 1965, p. 370-89.

say C-1, being affected at a point by another textual tradition, say ʿUṯmān or Ibn Masʿūd. Auto-contamination takes two forms: (a) assimilation of parallels and (b) assimilation of nearby terms.

(a) *Assimilation of Parallels.*[85] This refers to a scribe changing the text, making it more similar to a parallel passage in the work being transmitted (auto-contamination) or in another branch of the textual tradition (cross-contamination). It is the first of these, assimilation of parallels by way of auto-contamination, that is relevant here.

The Qurʾān is—and calls itself—"self-similar" (Kor 39, 23). It is full of repeated sentences and phrases that differ in one, two, or a few words. That is why even today memorizers routinely find themselves adding or substituting a word inadvertently if the added word appears in a *similar* sentence in a *different* verse. One's knowledge of other passages shapes one's memory of the verse at hand, generating substitutions, additions, and deletions that hark to the parallel. In dictation, hearing and memory are shaped by associations formed by previous exposure to the Qurʾān. While assimilation of parallels does occur in written transmission, it is more common in dictation because of the greater role of memory.[86]

An omission or minor change does not require any additional explanation: it can occur naturally as a first-tier transformation. If the variant is *also* a candidate for auto-contamination, then its occurrence becomes even more understandable. But additions of major elements are different. They do not occur as easily as omissions. They cannot be considered as first-tier unless a first-tier mechanism, such as the assimilation of parallels, can be invoked to explain them.

(b) *Assimilation of Nearby Terms.* It is more likely for a word to be used by mistake at a certain point if it is used in a nearby passage. A word is on the scribe's mind if he heard it a moment ago or if he expects to hear it soon due to prior familiarity with the passage at hand. Such a word can insinuate itself into the writing.[87]

---

[85] Assimilation of parallels, also called "harmonization", is a very common source of change in New Testament manuscripts. See Burnett Streeter, *The Four Gospels*, p. 139-44; E.C. Colwell, "Scribal Habits", p. 377-8; Metzger, *Text*, p. 193, 197-9; Peter M. Head, "Observations on Early Papyri of the Synoptic Gospels, Especially on the 'Scribal Habits,'" *Biblica*, 71 (1990), p. 240-7; Head, "Habits".

[86] If the scribe taking dictation subsequently compares the result with the written source and then makes corrections, the results will look different: there will be far fewer changes than in dictation alone. The extent to which this relatively careful variety of dictation may have played a role in the transmission of the ʿUṯmānic Qurʾān is worth investigating.

[87] Assimilation of nearby terms, or "harmonization to the immediate context", is an important category of scribal error in New Testament manuscripts. See Colwell, "Scribal Habits", p. 377-8; Royse, "Scribal Tendencies", p. 246; Head, "Observations"; Head, "Habits".

4. *Phonetic Conservation in Major Substitutions.* Because of the aural dimension, if a major element changes, more often than not it will *sound* like the original. For example, the same words may appear in a different order, or a verb may be changed to another verb that has the same root. Some types of phonetic conservation are listed in the left column of Table 6:

Table 6. *Types of Phonetic Conservation of Major Elements, with Examples.*

Type of Difference	Verse	ʿUṯmān	C-1
word order (transposition)	5, 41	*wa-lahum fī l-āḫirati*	*wa-fī l-āḫirati lahum*
another verb, same root	2, 196	*istaysara*	*tayassara*
verb conjugated for different person	15, 65	*tuʾmarūna*	*yaʾmurūna*
different form of plural	5, 43	*al-nabiyyūna*	*al-anbiyāʾu*
singular vs. plural	2, 219	*al-āyāti*	*al-āyati*
verb tense or mood	5, 52	*fa-yuṣbiḥū*	*fa-yuṣbiḥūna*
active vs. passive	5, 47	*anzala llāhu fīhi*	*unzila ʿalayhim fīhi*

Of the 25 cases of major substitution, *i.e.* cases where C-1 and ʿUṯmān use different words or phrases, 18 involve similar-sounding variants, while only 7 involve variants that do not sound similar. Appendix 1 lists the differences (in major elements) between C-1 and ʿUṯmān that are characterized by phonetic conservation, while Appendix 2 (c) lists those that are not.

5. *Common or Frequent Terms.* First-tier transmission generates no or very few *additions* that cannot be accounted for by auto-contamination. If such an addition does occur, however, it will have to be a frequently-used word. In any case, first-tier changes would *not* include the addition of a word that is unexpected *and* uncommon.

First-tier transmission generates, likewise, few *substitutions* that cannot be accounted for by either phonetic conservation or auto-contamination. Yet such substitutions can occur more easily than additions, since it is easier to misremember a word than "remember" a word that has not been said. If this happens, memory usually substitutes a familiar word with another word that is familiar or frequent, or a frequent word with another frequent word. A word is "familiar" if it is used elsewhere in the Qurʾān. (Less commonly does memory replace a rare word with a frequent word, and even less commonly does it replace a familiar word with a rare word. These last two transformations are *not* first-tier.)

With less exactness in transmission, *e.g.* due to faster dictation or purely oral transmission, some of the first-tier tendencies are progressively relaxed. It is convenient to refer to the less precise forms of transmission as "higher-tier". Higher-tier transmission practices still possess the hallmarks of dictation listed above, but they also show additional types of changes. For example, phonetic conservation is obeyed less frequently.

The most important observation one can make concerning C-1 is that its differences with the standard Qur'ān generally do not exceed first-tier variations. This confirms that the appropriate model is one of transmission by dictation rather than pure orality. Dictation explains not only why C-1 and 'Uṯmān are generally the same, even when it comes to minor elements, but also how the differences may have arisen. In particular, the role of aurality is evident in the large proportion of variations in minor elements, phonetic conservation of major elements, and the significant role of auto-contamination.

To say that the differences between 'Uṯmān and C-1 are first-tier begs the question of directionality. Some of the five tendencies listed above are asymmetric, as they can accommodate change in one direction more easily than the other. These can be used to see if it is easier to get, by dictation, from codex A to codex B or the other way around. I will show that going from 'Uṯmān to C-1 would be first-tier, but not the other way around. This is the second of two methods applied here to answer the question of priority.

The first method applied here, stemmatics, constructs a stemma, *i.e.* a family tree, using patterns of agreement and disagreement among text types. To do so, one needs to compare at least three texts. Therefore, I bring into play the codex of Ibn Mas'ūd, a non-'Uṯmānic text which, while not attested in manuscripts, is described in literary sources. Comparisons will be made with other Companion codices where data is available. This exercise presupposes that there is a significant kernel of truth in what the literary sources say about the codex of Ibn Mas'ūd. This is a point that requires discussion before proceeding further.

The Codex of Ibn Mas'ūd and other non-'Uṯmānic Qur'āns

Literary sources speak of codices of Companions of the Prophet such as Ibn Mas'ūd (popular in Kūfa),[88] Ubayy b. Ka'b (popular in Syria and Baṣra),[89] Abū Mūsā l-Aš'arī (Baṣra and Yemen),[90] and others. The codex of 'Abd Allāh

[88] For a discussion of the *muṣḥaf* of Ibn Mas'ūd, see Rāmyār, *Tārīḫ-i Qur'ān*, p. 353-64.
[89] For a discussion of the *muṣḥaf* of Ubayy b. Ka'b, see Rāmyār, *Tārīḫ-i Qur'ān*, p. 340-53.
[90] For a discussion of the *muṣḥaf* of Abū Mūsā l-Aš'arī, see Rāmyār, *Tārīḫ-i Qur'ān*, p. 380-4.

b. Mas'ūd, uniquely among the Companion codices, reportedly did not have the last two *sūras* of the standard Qur'ān, which together amount to forty-three words, which happens to be exactly the number of words found in this sentence.[91] In the codices of Ibn Mas'ūd and Ubayy b. Ka'b, the *sūras* are arranged in different orders relative to one another; but, within the *sūras*, the codices generally have the same arrangement of sentences as the standard Qur'ān. There were, however, occasional differences at the level of words and phrases. These differences have been reported in a variety of Muslim traditions from the first two centuries AH that are preserved in later Muslim literary sources. It would be naive to accept all such reports as reliable descriptions of Companion codices; but C-1 proves it equally unwise to dismiss them all as unreliable.

Among the reports about Companion codices, one tradition stands out because of its methodical nature. It is a report of the prominent second-century Kūfan traditionist and Qur'ān-reciter, Sulaymān b. Mihrān al-A'maš (d. 147/765, Kūfa). He lists Ibn Mas'ūd's "reading" at over 150 points.[92] The list includes many single-letter variants, such as the use of *wa-* ("and") versus *fa-* (punctuation). It also includes more important variations, such as substitutions, transpositions, additions, and omissions of words, although most of these make little difference to the meaning. The routine inclusion of minute variants, combined with the systematic nature of the list, creates the impression that the bulk of it is based on an actual codex. Furthermore, on the face of it very few variants appear to reflect a later development or to have been generated to further a theological or sectarian agenda.[93]

---

[91] It is also said that the codex of Ibn Mas'ūd lacked the first *sūra* (which is twenty-five words long). This may have been the case. But al-A'ẓamī says that this is contradicted by the fact that Ibn Mas'ūd's reading includes variants for this *sūra*. He also points out that Ibn al-Nadīm saw a *muṣḥaf* ascribed to Ibn Mas'ūd that had the *sūra* (al-A'ẓamī, *The History of the Qur'ānic Text*, p. 235-6; cf. Ibn al-Nadīm, *al-Fihrist*, p. 29, where he says that the manuscript was from AH 200). Note that al-A'maš, at least, does not report variants of Ibn Mas'ūd for the first *sūra*. For references to the primary sources with these reports about the first and the last *sūras*, see Modarressi, "Early Debates", p. 13, footnote 48; al-A'ẓamī, *The History of the Qur'ānic Text*, p. 236. Al-A'ẓamī dismisses the reports he quotes as self-contradictory (*ibid.*); but a plausible harmonization would not be difficult to imagine. In the final analysis, however, it must be admitted that the reports are uncertain.

[92] Ibn Abī Dāwūd, *al-Maṣāḥif*, p. 57-73.

[93] One variant of Ibn Mas'ūd as reported by al-A'maš gives me pause: in Kor 3, 7, in lieu of the standard phrase, it reads *wa-in ḥaqīqatu ta'wīlihi illā 'inda llāhi wa-l-rāsiḫūna fī l-'ilmi yaqūlūna āmannā bihi* (Ibn Abī Dāwūd, *al-Maṣāḥif*, p. 59). It is possible that this is a deliberate change intended to remove the ambiguity of the standard version. A critic might object that this is a conspiracy theory. But it is not conspiracy-minded to allow for the possibility of deliberate change in an individual report (which may or may not have been accepted by others) in the way it

Among the variants ascribed to Ibn Mas'ūd and Ubayy b. Ka'b, several are anomalous: they add large new clauses or phrases absent in the 'Uṯmānic text that are not explainable in terms of the aural dimension of dictation—and these happen to explicitly endorse the 'Alīd cause.[94] It would be natural to dismiss these as tendentious interpolations; but to generalize this negative judgment to the totality of reported variants would be to throw the baby out with the bathwater. Rather, one expects some of the sources of non-'Uṯmānic readings to be more trustworthy than others. In particular, as explained in the last paragraph, the Ibn Mas'ūd variants given by al-A'maš do not seem to be tendentious. Nevertheless, it remains true that al-A'maš does not disclose his sources,[95] and research clearly is needed into the sources and transmission of his and other non-'Uṯmānic variants.

A key point about al-A'maš's version of Ibn Mas'ūd's codex is that more often than not its variants are related to the 'Uṯmānic text through first-tier transformations. Minor elements form a large category, and, as shown in Table 7, the use of synonyms and the tendency to conserve major elements phonetically are in evidence. There is also evidence of auto-contamination, most noticeably in the form of assimilation of nearby terms.

Comparing Table 7 (on Ibn Mas'ūd) with Table 6 (on C-1) or Appendix 1 reveals striking similarities. In many cases, Ibn Mas'ūd (known via al-A'maš) relates to 'Uṯmān in ways that are familiar from the ways in which C-1 relates to 'Uṯmān. That lends credence to the historicity of the Ibn Mas'ūd codex and suggests that it indeed possessed many of the variants ascribed to it. By this, I do not mean that C-1 and Ibn Mas'ūd share actual variants, which they do in a few cases, but that they share some of the same *types* of variants. In general, every type of variant found in C-1 is found also in Ibn Mas'ūd. However, Ibn Mas'ūd also has some higher-tier types not found in C-1. (For these types, one cannot yet claim corroboration from manuscripts, warranting some caution regarding their authenticity.) This means that, on the whole, the standard Qur'ān is closer to C-1 than it is to the codex of Ibn Mas'ūd—qualitatively that is, but not in terms of the quantity of variations.

It is not just by way of verbal variants that C-1 supports the historicity of non-'Uṯmānic codices and the reliability of some of the reports about them; there are also the similarities in the order in which some *sūra*s appear in C-1 and the *sūra* sequences reported for the codices of Ibn Mas'ūd and Ubayy b. Ka'b, as shown in Table 8.

---

would be to speak of deliberate changes accepted by entire communities in an already well-known divine text.

[94] For the details, see Modarressi, "Early Debates", p. 25-6.

[95] Al-A'ẓamī, *The History of the Qur'ānic Text*, p. 234.

Table 7. ***Examples of the phonetic conservation of major elements between Ibn Masʿūd and ʿUṯmān. Compare to Table 6.***

Type of difference	Verse	ʿUṯmān	Ibn Masʿūd (via al-Aʿmaš)
word order (transposition)	3, 156	*wa-llāhu bi-mā taʿmalūna baṣīrun*	*wa-llāhu baṣīrun bi-mā taʿmalūna*
synonymous verb, same root	2, 202	*kasabū*	*iktasabū*
verb conjugated for different person	2, 83	*taʿbudūna*	*yaʿbudūna*
different form of plural	13, 42	*l-kuffāru*	*l-kāfirūna*
singular vs. plural	34, 37	*l-ġurufāti*	*l-ġurfati*
tense or mood	17, 44	*tusabbiḥu*	*sabbaḥat*
active vs. passive	32, 17	*qāla aʿlamu*	*qīla aʿlamu*

Table 8. ***The* sūra *order in C-1 compared to those reported for Ibn Masʿūd and Ubayy b. Kaʿb. The numbers given represent the ʿUṯmānic order.***

C-1	Ibn Masʿūd	Ubayy b. Kaʿb
63, 62, 89, 90	The *Fihrist* list:[96]	The *Fihrist* list:[98]
	63, 62,... [twenty-nine intervening *sūras*], 89, 85, 84, 96, 90	63, 62, 65, 89,... *sūra* 90 does not appear on the list.
	The *Itqān* list, quoting al-Aʿmaš:[97]	The *Itqān* list:[99]
	63, 62,... [twenty-seven intervening *sūras*], 89, 85, 84, 96, 90	63, 62, 66, 89, 90

[96] Ibn al-Nadīm, *al-Fihrist*, p. 29; Rāmyār, *Tārīḫ-i Qurʾān*, p. 356-9; Neal Robinson, *Discovering the Qur'an: A Contemporary Approach to a Veiled Text*, 2nd ed., London, SCM Press, 1996, p. 263-6; Arthur Jeffery, *Materials for the History of the Text of the Qurʾān: The Old Codices*, Leiden, Brill, 1937, p. 22.

[97] Al-Suyūṭī, *al-Itqān fī ʿulūm al-Qurʾān*, n.p., Dār al-fikr, 1416/1996, I, p. 176; Rāmyār, *Tārīḫ-i Qurʾān*, p. 359-61; Robinson, *Discovering the Qurʾān*, p. 263-6; Jeffery, *Materials*, p. 23.

[98] Ibn al-Nadīm, *al-Fihrist*, p. 29-30; Rāmyār, *Tārīḫ-i Qurʾān*, p. 344-5; Jeffery, *Materials*, p. 115.

[99] Al-Suyūṭī, *al-Itqān*, I, p. 175-6; Rāmyār, *Tārīḫ-i Qurʾān*, p. 346; Jeffery, *Materials*, p. 115.

One may make two observations about C-1's *sūra* ordering. First, by virtue of the 63-62 sequence, it can be considered closer to the Ibn Masʿūd and Ubayy b. Kaʿb codices than to ʿUṯmān. Second, it is over-all much nearer to Ubayy b. Kaʿb than it is to Ibn Masʿūd.

Which Text Type is Older? Method I: Stemmatics

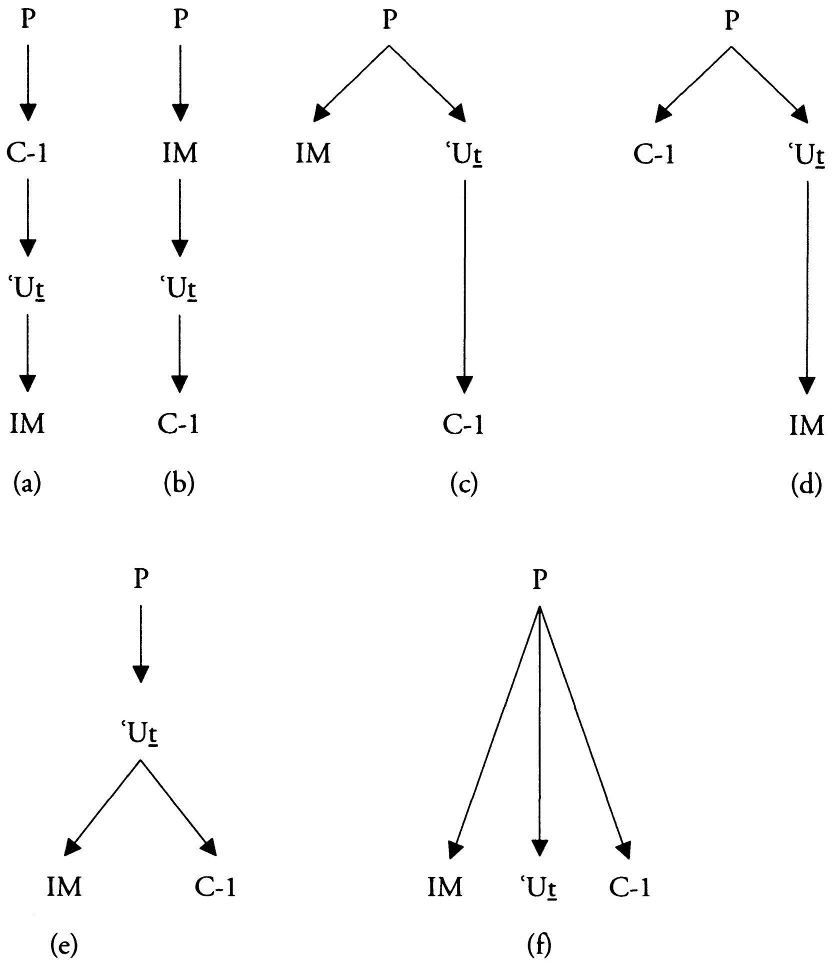
I shall now attempt to construct the family tree, or stemma, of three *text types*, those of Ibn Masʿūd, ʿUṯmān, and C-1, based on the patterns of agreement and disagreement among them. It is impressive that Ibn Masʿūd's codex agrees with C-1 several times, as shown in Table 4; but far more striking is how exceptional such agreements are. Usually, where there is a difference between ʿUṯmān and Ibn Masʿūd (or, for that matter, between ʿUṯmān and any other Companion codex), C-1 sides with ʿUṯmān. In other words, in almost every case in which there is a disagreement between the three, ʿUṯmān is in the majority: either it is ʿUṯmān and C-1 against Ibn Masʿūd, or it is ʿUṯmān and Ibn Masʿūd against C-1. ʿUṯmān is rarely alone: that is, seldom do Ibn Masʿūd and C-1 agree against ʿUṯmān. Therefore, ʿUṯmān occupies a central position, which may be represented as follows:

C-1 — ʿUṯmān — Ibn Masʿūd

Depending on which of its three nodes one holds up or pulls down, this pattern yields different stemmata. The constraint imposed by this pattern cuts down the number of feasible stemmata of the text types including a prototype from sixteen to the six shown in the following diagram, labeled (a) through (f). The case in which ʿUṯmān is pulled down and the others are held up is shown later in (g).[100]

Here, P stands for the "Prototype"—as I shall argue later, to be identified with the Prophet. IM, ʿUṯ, and C-1 refer respectively to the text types (or textual traditions) of Ibn Masʿūd, ʿUṯmān, and C-1. If one does not posit a prototype, one obtains only stemmata (a), (b), and (e), with P removed. This set of stemmata is what the verbal features yield if the only information that is taken into account is the pattern of verbal agreement and disagreement between the codices and if the exceptional cases are set aside. The question is whether there are other grounds for preferring one stemma over another. The following paragraphs show that there are reasons to favor stemmata (e) and

[100] On the construction of stemmas, see L.D. Reynolds and N.G. Wilson, *Scribes & Scholars: A Guide to the Transmission of Greek & Latin Literature*, 3rd ed., Oxford, Clarendon Press, 1991, p. 211-3. I owe this reference to Michael Cook.



(f), the latter being more probable. Note that under (f), the fact that ʿUṯ is in the majority entails that it is more reliable than IM and C-1.

All the stemmata would have to describe acts of transmission *before* the promulgation of the ʿUṯmānic Qurʾān. For by the time ʿUṯmān had Qurʾāns copied and sent to the cities, copies conformed very closely to the wording of their sources, differing from them typically only in units of one or two letters;[101] and, thereafter, as manuscript evidence shows, changes were of the

[101] Cook, "The Stemma of the Regional Codices of the Koran", p. 90-1. For a discussion, see the section above on the upper text.

limited type expected of written transmission. The post-ʿUṯmān range of verbal variations was thus relatively small. While this observation does not point to a stemma, it highlights a crucial fact that helps narrow down the choices: there was a drastic transition from a mode of transmission generating first-tier changes due to semi-orality to a mode of transmission with relatively small changes that are characteristic of copying from manuscript. This radical shift took place in the twenty years or so before ʿUṯmān's standardization, not a long interval. Furthermore, it took place either at a specific point in time (*e.g.* as soon as P had dictated the text in scenario (f)) or gradually. Now, if one were to accept (a), (b), (c), or (d), that would make the variants the product of two or three successive generations of textual reproduction, and one would expect the abovementioned shift (in reproductive fidelity) to be reflected in the codices by virtue of the intermediate codex, ʿUṯ, being verbally *much* closer to the youngest codex than the oldest.[102] The data does not quite fit such a pattern. ʿUṯ is slightly closer to C-1 in terms of the quality of variants but closer to IM in terms of their quantity. It also counts against these stemmata that IM and C-1 occasionally agree against ʿUṯ. These considerations leave (f) as the plausible choice, with (e) being the second best. The stemma (g), to be discussed later below, remains as plausible as (f).

The fact that fidelity in transmission increased rapidly within a couple of decades shows that the biggest changes must have been introduced in the first stage of transmission. (This reminds one of other situations where the largest variations are the oldest, such as in the *Ḥadīṯ* literature or the transmission of New Testament manuscripts.)[103]

The choice (f) would also follow from a scenario in which the text was originally dictated, and thereafter it was copied from manuscripts or was dictated and then corrected against the written source. On this scenario, the differences that are due to orality must go back to the original moment of

[102] If either (a) or (b) were the true stemma, then the wording of either C-1 or IM would be much more similar to ʿUṯ than to the other. If IM were the one much more dissimilar to ʿUṯ, we would prefer (b) to (a); but if C-1 were the one much more different, we would prefer (a) to (b).

[103] Spencer *et al.* lend credence to the "ideas that most variants arose early in the history of the Greek New Testament,... and that later copies introduced fewer variants" (Matthew Spencer, Klaus Wachtel, and Christopher Howe, "Representing Multiple Pathways of Textual Flow in the Greek Manuscripts of the Letter of James Using Reduced Median Networks", *Computers and the Humanities*, 38 (2004), p. 1-14). This comparison, however, should not blind one to the profound differences between the New Testament and the Qur'ān, including the fact that the Qur'ān was already scripture at its inception and constantly identifies itself as such, and the fact that Islam became a state religion already as the Qur'ān was being revealed. These factors explain why the text of the Qur'ān was stabilized so much more rapidly than that of the New Testament.

dictation, and thus only (f) fits the data.[104] The Qur'ān perhaps lends some support to this scenario by its self-description as a written text akin to other divine scriptures. It identifies itself as having been divinely revealed to the Prophet and as having taken the form of an oral utterance by him (Kor 75, 16-8; 98, 2-3) that was put in writing (Kor 25, 5; 98, 2-3),[105] revelations that Muslims were expected to recite as frequently as possible (Kor 73, 20), a scripture akin to other holy books, the Torah and the Gospels (Kor 9, 111). One may read these Qur'ānic statements as being consistent with the idea that the Prophet dictated the Qur'ān to a number of scribes and that it was thereafter transmitted in writing, like the Torah and the Gospels. Note, however, that even without this scenario stemma (f) remains plausible.

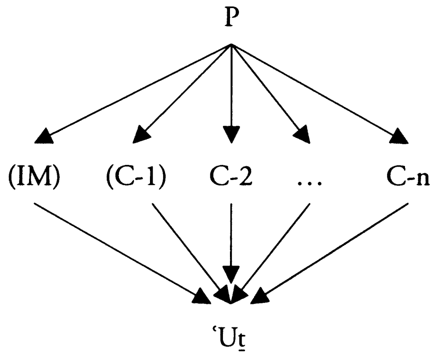
In short, the options are (e) and (f)—stemma (f) being more likely.[106] The former (e) would make ʿUṯmān the common source of Ibn Masʿūd and C-1. The latter (f) would make the three text types independent reproductions of

[104] The scenario in which the Qur'ān was dictated by the Prophet and thereafter copied by written transmission would be plausible if there were no interdiction on written transmission. And indeed there is no sign of any such interdiction for the scripture. On the other hand, attitudes toward the writing of non-scriptural religious literature were ambivalent. For the state of this controversy in the second century AH, see Michael Cook, "The Opponents of the Writing of Tradition in Early Islam", *Arabica*, 44 (1997), p. 437-530.

[105] Rāmyār cites a number of verses to support the idea that the Qur'ān was put in writing by Muḥammad. The most definitive one is Kor 25, 5. There is also Kor 98, 2-3. The Qur'ān's mention of teaching by the pen (Kor 96, 1-5) may be an allusion to divine scriptures and the role of writing in their transmission. The verses at Kor 74, 52 and 6, 91 are highly suggestive. The cases of Kor 80, 11-6; 85, 21-2, and 56, 77-80 are relevant, although, as Rāmyār acknowledges, less conclusive. Rāmyār also gives a variety of other Qur'ānic references to writing, copying, and their accoutrements such as paper and parchment. He also highlights the Qur'ān's references to the Jewish Bible as a physical object. See Rāmyār, *Tārīḫ-i Qur'ān*, p. 257-8, 275-9. It would be of interest to know how Jewish scriptures were transmitted in Arabia on the eve of Islam. Cook has shown that some early Muslim ideas about written and oral scriptures resonated with Jewish notions (Cook, "Opponents", cited in the last footnote).

[106] One may also attempt to choose among the stemmata using the data of the *sūra* ordering of the codices. The fact that the orderings of Ibn Masʿūd and C-1 are closer to each other than they are to ʿUṯmān would exclude all stemmas except (e) and (f). If we went with (f), then one of the following three scenarios would explain the similarity of IM and C-1: (*i*) they reflected P's *sūra* ordering better than ʿUṯ, (*ii*) one influenced the other ("contamination"), or (*iii*) they independently used a similar principle to order the *sūra*s. If we went with (e), then *i* would become moot and we would have to choose between *ii* and *iii*.

Note that if *ii* or *iii* are true, however, then *all* the stemmata in the diagram would regain plausibility and *sūra* ordering would not help us choose between them. I place this discussion in a footnote because it is very easy to imagine *ii* and *iii* as valid scenarios for *sūra* ordering, especially if, as it seems to be the case, at the time of the Prophet many *sūra*s were free-standing, not yet rigidly ordered in a fixed sequence. Scenario *iii* is especially plausible since the *sūra* sequences ascribed to Ibn Masʿūd and Ubayy b. Kaʿb appear intended to arrange most *sūra*s in order of decreasing length (Robinson, *Discovering the Qur'ān*, p. 263-6).



(g)

the prototype; however, ʿUṯmān's version would be the most reliable of the three, for in cases of disagreement it is generally in the majority. Only exceptionally do Ibn Masʿūd and C-1 agree against ʿUṯmān. These commonalities are due to "cross-contamination" (one textual tradition influencing the other), or are independent "convergent adaptations", or indeed go back to the prototype. As discussed further below, in this last case the possibility remains that, as some Muslim traditions maintain, the Prophet had disseminated revised versions of the prototype.

The analysis so far has treated the ʿUṯmānic codex as a package in which different parts have the same textual history. If, however, one entertains the possibility that ʿUṯmān is a hybrid text based on a number of Companion codices, C-2,..., C-n, and possibly IM and C-1, then the stemma will be (g). In this scenario, the fact that ʿUṯmān's version gives the majority reading means that it emerged as part of an attempt to find consensus. Where two codices disagreed, ʿUṯmān's committee would have sought independent corroboration by other Companions. This happens to be more or less how a modern scholar would set out to reconstruct the prototype if he had access to a number of independent copies of a common source. And, indeed, such a procedure can be expected to have produced a better text than any of the individual codices (IM, C-1, C-2, etc.), *i.e.* one that better preserved P's wording. Stemma (g) and the ideas in this paragraph are due to Michael Cook.[107]

[107] Michael Cook, "Concluding Remarks", Colloquium on the Evidence for the Early History of the Qur'ān, Stanford University, July 31, 2009.

This scenario does not ascribe to the committee an unreasonably high degree of sophistication. The practice of verifying a report by seeking corroboration is a matter of common sense, and it is mentioned in the Qur'ān (Kor 24, 12-3; 2, 282). And while for the purpose of this essay I suspend judgment on the accuracy of reports of individuals, it is worthwhile noting that two reports, with independent *isnāds*, traceable to two different scribes in 'Uṯmān's team state that in cases of disagreement over wording, the disputed point was postponed or the spot was left blank so that it could be dealt with later. If true, that would support hybridity.[108]

To sum up, the evidence of the palimpsest inspires confidence that literary sources carry partly valid information about non-'Uṯmānic codices. In particular, I provisionally accept al-A'maš's description of the codex of Ibn Mas'ūd (above, p. 49).[109] With three codices available, it is natural to seek to determine genetic relationships. 'Uṯmān occupies a central position between the other two. This indicates either that 'Uṯmān is the more faithful copy of a common source, the prototype, or that it is a hybrid reflecting the majority reading of Companion codices. Chronological constraints help identify the prototype with the Prophet's own recension, and the Prophet as the person who dictated it—an argument that will be spelled out in a later section below.

Which Text Type is Older? Method II: Internal Evidence and Polarity Analysis

The application of stemmatics above relies on certain propositions: that we know Ibn Mas'ūd's codex and that the biggest verbal changes were introduced at an early stage in transmission. These may be excellent premises; but it would be gratifying if there were an additional argument that did not rely on them.

I turn to analyzing the internal evidence. Do 'Uṯmān and C-1 have the same character relative to each other, or is there an asymmetry that shows that one of them has the older wording? To answer this question, one should ask which is easier to imagine: scribal mistakes that transform 'Uṯmān's wording to that of C-1, or scribal mistakes that change C-1 to the 'Uṯmānic

---

[108] Ibn Abī Dāwūd, *al-Maṣāḥif*, p. 21-2 and 25-6; cf. Rāmyār, *Tārīḫ-i Qur'ān*, p. 426, where the author cites additional sources for one of these reports.

[109] Although I used the codex of Ibn Mas'ūd as reported by al-A'maš in attempting to construct the stemma, and although I have attempted to show that al-A'maš's report is largely credible, it is important to note that the validity of my argument is not sensitive to the details of the A'maš report. If instead of that report, I used *all* or *some* of the other reports about the codex of Ibn Mas'ūd, Ubayy b. Ka'b, or any other Companion, then 'Uṯmān would still be in the majority, and the stemmata would be the same.

version? In other words, is the transformation of one codex to the other more consistent with first-tier changes (p. 44-7) than the other way around? I will show that, indeed, the transformation of 'Uṯmān to C-1 can be explained in terms of first-tier changes, but the transformation of C-1 to 'Uṯmān cannot.[110]

This is a highly non-trivial observation. Suppose that the dictation of a common source generated two codices, A and B. Even if A and B were both first-tier copies of their common source, in general the relationship between A and B would not be reducible to the first tier, *i.e.* one could not show that one can get from A to B (or from B to A) with first-tier transformations. (For example, A could omit some words and B some other words, so each would have major terms lacking in the other which could not be explained by auto-contamination, etc., making it impossible to show that you can get from one codex to the other by first-tier transformations.) So, the fact that one *can* get from 'Uṯmān to C-1 by first-tier transformations means either that 'Uṯmān (*i.e.* its wording) is the source of C-1 or that it is an accurate version of their common source.

Let us now see if it is as easy to get from one codex to the other as the other way around. The most useful set of differences for the analysis are cases in which one codex has a major element in a verse that the other codex does not. I call these major pluses. Consider, first, those cases in which a passage in C-1 has a word not found in the same passage in 'Uṯmān. Let us label such extra terms as pluses of C-1. Similarly, there are cases in which a passage in 'Uṯmān has an extra term not found in C-1's version of that passage. These are the pluses of 'Uṯmān.

(1) *Omissions.* It is easier for a scribe taking dictation to drop and change major words he has heard than to add major words he has not heard. 'Uṯmān has fourteen major pluses while C-1 has ten (see Appendix 2). As for minor pluses, 'Uṯmān has twelve while C-1 has only four (see Appendix 1). The total is twenty-six pluses in 'Uṯmān versus fourteen in C-1. This suggests that through dictation it is easier to get from 'Uṯmān (or from a prototype that 'Uṯmān represents faithfully) to C-1 than the other way around. However, numbers alone are not decisive. It is the nature of the pluses that clinches the argument: it will be noted below (3) that all the major pluses of C-1 can be ascribed to other first-tier mechanisms such as auto-contamination, but that cannot be done for all the major pluses of 'Uṯmān.

(2) *Omissions in Lists.* Keeping several items in mind simultaneously can be a challenge. If a scribe who takes dictation forgets a word, it is easy to see

[110] In this entire section, by "transforming codex A to codex B", I mean "getting from *the wording* of codex A to *the wording* of codex B".

how this could happen if the word is an item on a list. It is, therefore, significant that of the pluses of ʿUṯmān, two appear in lists of three items apiece. In C-1, these lists appear with only two items. For the details, see Appendix 2, section (a) ("Omissions"), Kor 2, 196 and Kor 2, 218. It is easier to see how one could get from ʿUṯmān (or its prototype) to C-1 than the other way around.

(3) *Major Pluses.* The most important pattern is the way in which the pluses in C-1 differ from those in ʿUṯmān. I begin with C-1. There are ten major pluses in C-1, and these are listed and annotated in Appendix 2, section (b) ("Additions"). Prior to analysis, these words can be considered either additions of C-1 or omissions of ʿUṯmān.

Their most telling attribute is that every one of them, without exception, appears in a parallel passage elsewhere in the standard Qurʾān that *has* that word. (A passage is a "parallel" if it, too, contains some of the words or phrases surrounding the plus.) Take Kor 2, 193. Here, C-1 has "And fight them until persecution is no more and religion is *all* for God." The word "all" is a plus of C-1; ʿUṯmān does not have it. Yet ʿUṯmān has this exact sentence, including "all", in another verse, namely Kor 8, 39. The extra "all" of C-1 at Kor 2, 193, therefore, is a candidate for assimilation of the parallel found in Kor 8, 39. That is to say, the word "all" might have occurred to C-1 because of the association with the other verse.

It is easy to see how the nine major pluses of C-1 could be secondary to the ʿUṯmānic version: they are candidates for auto-contamination. But can one make the same argument, in reverse, for the pluses of ʿUṯmān, to show that ʿUṯmān could be secondary to C-1? Is the situation symmetric? No, it is not.

It is different with the pluses of ʿUṯmān. There are fourteen major pluses in ʿUṯmān—see Appendix 2, section (a) ("Omissions"). Among them are five that occur in sentences that are not attested or paralleled anywhere else in the Qurʾān. Nor do these five terms show up anywhere else with phrases that surround them in the verse at hand. Therefore, they could not have been generated by auto-contamination. These "orphans" are as follows: Kor 2, 196 (*aw ṣadaqatin*), 2, 217 (*wa-kufrun bihi*), 2, 222 (*fa-ʿtazilū*), Kor 5, 42 (*fa-in ǧāʾūka*), and 63, 1 (*qālū*); although one perhaps should discard the last two cases—Kor 5, 42 because of its proximity to *iḏā ǧāʾaka* in Kor 5, 48 (making it a candidate for assimilation of nearby terms), and Kor 63, 1 because *qālū* is a very common term in the Qurʾān.

It is worth giving a systematic and precise paraphrase of the argument. First consider the pluses of C-1. If the wording of C-1 is older than that of ʿUtmān, this would mean that ʿUtmān omitted these terms. Omissions are

consistent with first-tier transmission. So, it would be easy to get from C-1 to ʿUṯmān through dictation. On the other hand, if the wording of ʿUṯmān is older that that of C-1, the pluses of C-1 would be items added by C-1. Additions cannot be first-tier changes unless they are candidates for auto-contamination. Indeed, all the pluses of C-1 can be explained as possible cases of auto-contamination: these terms, without exception, are found in parallel passages. Thus, it is equally easy to imagine ʿUṯmān as the older text. In sum, the pluses of C-1 do not favor one codex over the other. This is because depending on which text is older, the pluses of C-1 are either omissions of ʿUṯmān or auto-contaminations of C-1.

The situation is different with the pluses of ʿUṯmān. If ʿUṯmān has the older wording, these would be omissions of C-1. Omissions are first-tier changes, and this makes it easy to imagine getting from ʿUṯmān to C-1. On the other hand, if C-1 is older, these pluses would be additions of ʿUṯmān. But additions cannot be first-tier changes unless they are candidates for auto-contamination or some other first-tier process. In fact, several pluses of ʿUṯmān cannot be due to auto-contamination, as they are not attested elsewhere in similar contexts. Nor does any other first-tier process explain them. Therefore, it is difficult to imagine getting from C-1 to ʿUṯmān through dictation. In sum, some pluses of ʿUṯmān can be omissions of C-1 but not auto-contaminations of ʿUṯmān, a situation consistent with the priority of ʿUṯmān.

The conclusion is that ʿUṯmān has the older wording. The pluses of C-1 are due to assimilation of parallels by C-1, while the pluses of ʿUṯmān are omissions by C-1. Some of these omissions of C-1 may have involved (and been facilitated by) assimilation of parallels.

(4) *Assimilation of Nearby Terms.* One's memory is also shaped by the things one has heard most recently or, if one is familiar with a text, by what one has previously heard in the vicinity of the point at hand. For this reason, it is necessary to examine nearby passages to check for the possibility of influences from nearby terms and parallels. In three instances, ʿUṯmān's phrasing (where it differs from C-1) parallels nearby passages: Kor 2, 217 (*ʿan dīnikum*), Kor 2, 221 (*bi-iḏnihi*), Kor 5, 42 (*fa-in ǧāʾūka*). On the other hand, there are five cases in which the variants in C-1 parallel nearby passages: Kor 5, 44 (*yaḥkumūna bi-mā nazzala llāhu fīhā* and *yaḥkumu bihi*), 5, 45 (*ʿalā banī Isrāʾīla*), 63, 3 (*ḏālika bi-annahum qawmun*), possibly Kor 63, 11 (*li-ġadin*), possibly Kor 5, 45 (*li-l-muttaqīna* vs. *li-qawmin yūqinūna*, provided the last word is *yūqinūna* and not *yuʾminūna*).

The greater number of *nearby* parallels in the case of C-1 might suggest that C-1 is secondary. However, caution is in order, as the tally of three

against five does not represent a large disparity, and it could be due to chance. Clearly, it would be desirable to study the other extant folios of the manuscript to determine whether a trend emerges.

(5) *Substitutions.* There are eight cases in which each codex uses a different major word or phrase that cannot be classified under phonetic conservation. These "substitutions" are listed in Appendix 2, section (c). In these cases, the two codices generally use synonyms or similar terms both of which appear elsewhere in similar contexts, or both of which are common terms; however, there are a couple of exceptions. In the case of Kor 5, 46 (*li-l-muttaqīna* vs. *li-qawmin yu'minūna* or *yūqinūna*), it is the version of 'Uṯmān that has significant parallels elsewhere. Normally, this would militate against 'Uṯmān's priority. (However, if one reads C-1's variant as *yūqinūna* instead of *yu'minūna*, then it becomes a candidate for assimilation of a nearby term, consistent with 'Uṯmān's priority). On the other hand, the case of Kor 63, 4 (*ra'aytahum* vs. *ǧā'ūka*) leans in the direction of 'Uṯmān's priority. But, there is no trend. Now, one may also consider the cases in which two different major terms that sound similar are used. These are included under the rubric "phonetic conservation" in Appendix 1, section (b). These do not evince a trend.[111]

The analysis is complete. The reader will note that I did not call any of the variants a deliberate change. With a little bit of imagination, anyone might concoct a theological or political motive for any major variant. (Given enough zeal for invoking religiously or politically-motivated changes, one can posit such motives even for variants that turn out to be imaginary.)[112] Yet such explanations are gratuitous in the case at hand, since a single principle gives a natural explanation for the data: everything is accounted for in a plausible way as a first-tier by-product of semi-orality. Given a falsifiable theory that explains the data in a simple way, it will be a violation of explanatory unification and simplicity to use that theory *plus* another one to explain the data.[113] The hypothesis of deliberate change, unlike the theory of first-tier change, fails to explain many things. Take, for example, the pluses of C-1. Say, for the sake of argument, that the scribe of 'Uṯmān's codex deliberately removed the word "all" in Kor 2, 193 because he did not like it for some theological or political reason. If so, why would he let "all" remain when the *exact* same sentence occurs in Kor 8, 39? Deliberate change fails to explain

[111] Another consideration, besides substitutions, that does not reveal a trend is a word's frequency in the Qur'ān. In both instances, it would be of interest to see whether a trend emerges in the larger sample of all extant folios of the manuscript.

[112] See above p. 21 and footnote 32.

[113] On explanatory unification, see Philip Kitcher, "Explanatory Unification", *Philosophy of Science*, 48/4 (1981), p. 507-31.

why C-1 does not add any major element that is not attested elsewhere.Deliberate change fails to explain why there are so many minor variations. It does not explain the use of synonyms. It also fails to explain why most substitutions obey phonetic conservation.

Lest the appeal to explanatory unification be an overly philosophical way to put the matter for some readers, it is useful to quote P. Kyle McCarter's commonsense formulation of the point:[114]

> *Look first for the unconscious error.* The better solution to a textual problem—when one solution assumes a textual accident and another an intentional alteration—is the first... Textual criticism is a rational activity, and there is a temptation to rationalize the changes that occurred in the text. The fact is, however, that deliberate alterations of the text were rare in antiquity, and textual accidents were common. The critic who thinks in terms of inadvertent errors, therefore, will do the more effective job of restoring the text.

In conclusion, there are several significant cases of asymmetry between the codices. *These disparities are precisely those one would expect to find if the textual differences were due to the aural dimension of taking dictation and if 'Uṯmān's text were earlier than C-1 in the sense of being closer to the original (or being the original).* This conclusion is falsifiable in principle, meaning that the method employed here would have objectively led us to the opposite conclusion if the trends had been the opposite of those observed. The method is not loose enough to lend itself to justifying any conclusion one desires. This fact has the following direct practical consequence. Most of the extant folios remain as yet unanalyzed. In them, one can expect to find exceptions to the trends identified here. If these exceptions are numerous enough to destroy the trends, then my conclusions will be undermined. If they are numerous enough to reverse the trends, then my conclusions will be refuted and reversed. If, however, the trends remain, then that will be an even more striking corroboration of 'Uṯmānic priority.[115]

[114] Peter Kyle McCarter, *Textual Criticism: Recovering the Text of the Hebrew Bible*, Philadelphia, Fortress Press, 1986, p. 23.

[115] To note a caveat, one should also be prepared for the possibility that different *sūra*s have different transmission histories due to the involvement of different scribes. It could be that a Companion who collected the Qur'ān did not hear every single *sūra* directly from the Prophet. It could be that he heard some *sūra*s personally, but relied on other Companions for the *sūras* he had not heard himself. This means different *sūra*s in C-1 could have different transmission histories before they were collected in one codex. The possibility (as maintained in Muslim traditions) that the Prophet himself disseminated different versions of the *sūra*s or (as indicated in the Qur'ān) that he revised the *sūra*s, if real, could be expected to make the data more complex. It may become necessary to consider these potential complications after the remaining folios have been studied.

The analysis conducted in this section captures a pattern of asymmetry in the data that shows that the two text types have different characters. This asymmetry yields a broad conclusion about the relative priority of one text type over the other. Such an analysis of trends or tendencies, however useful it may be for arriving at a conclusion about a codex as a whole, is generally not suitable for judging an individual word or phrase, just as, from the statistical fact that *on average* one year is warmer than another one, cannot infer anything with high probability about individual days.

Which Text Type is Older? A Third Method? Obvious Mistakes.

Devin Stewart has made a preliminary attempt to lay out the logic of textual emendation in the context of copyists' errors.[116] The method, for him, turns on probability—a probable word that is graphically similar to the given, improbable text is a candidate for being the original word—and probability may be assessed, among other ways, by frequency of occurrence: a word that is well-attested in the Qur'ān is more probable than one that is not.[117] He thus models textual corruption as a movement from a probable word to an improbable one.

The validity of Stewart's model for its intended context of written transmission is open to debate; but let us set that question aside and ask whether the model is applicable to a context for which he did not intend it: namely, scenarios in which memory plays a more prominent role, as in the case of a dictated text. In such a situation, it is not obvious that corruption would tend to run towards greater improbability, as measured, for example, by attestations elsewhere in the book. One would expect that in most cases the opposite would be the case.

Movement toward greater probability would be more common; but it cannot be denied that Stewart's model may generate exceptions, and this is one reason (among others) why the method employed in the last section looks for statistically significant trends rather than for fixed and inviolable patterns. In fact, one way to identify candidate instantiations of Stewart's model is to focus on the exceptions to the general trend. But then this method is not an independent way of distinguishing the secondary text type from the primary

[116] Devin Stewart, "Notes on Emendations of the Qur'ān", in *The Qur'ān in its Historical Context*, ed. Gabriel S. Reynolds, London, Routledge, 2008, p. 225-48.

[117] The link between probability and attestation is shown by Stewart's comment on Kor 56, 28-31: "It seems odd that bananas are mentioned here, for they do not occur elsewhere in the Qur'ān" (*Ibid.*, p. 233).

one, for it presupposes that one already has made that determination by identifying the trend.

Is all hope lost for using Stewart's approach or a variant thereof? There is a class of improbable words for which the direction of change is not in doubt, namely *extreme* improbabilities that are *prima facie* errors, and therefore may be classified as secondary without hesitation. These include obvious errors of the hand and obvious cases of mishearing or false memory. Both types occur in C-1. Errors of the hand (see Appendix 2, section (d)) are almost certainly singular readings that do not represent the text type and, therefore, imply little about the relative dates of the two text types; but obvious cases of mishearing or false memory might speak to the degree of accuracy of the codex as a whole. Instances of the latter in C-1 include *mā lahum* instead of *mā lahu* in Kor 2, 200 and *fa-iḫwānuhum* instead of *fa-iḫwānukum* in Kor 2, 220. (Here, I set aside cases of awkwardness in C-1, such as the use of *fa-yaqūlu rabbanā... aḫḫirnā... fa-aṣṣaddaqa wa-akun* in Kor 63, 10, the use of the singular *al-āya* in Kor 2, 219, the use of *bihā* instead of *bihi* in Kor 5, 45, and the use of *ḏālika bi-annahum qawmun lā yafqahūna* in Kor 63, 3, which results in the awkward repetition of *ḏālika bi-annahum.*)

IV. Conclusion

Implications for the Role of the Prophet

Fidelity in transmission increased rapidly in the early years of Islam. This proposition is corroborated well when one contrasts the relatively minor differences in the ʿUṯmānic branch, representing the limited caliber of the changes that could arise after about AD 650, with the relatively significant differences between ʿUṯmān, C-1, and Ibn Masʿūd. Within the ʿUṯmānic textual tradition, the changes are of the type one would associate with copying from manuscript or dictating and then correcting against a manuscript. By contrast, the relatively larger differences with C-1 are due to a measure of orality. The splitting of the branches took place before the internal developments in either branch. The branching point represents the prototype, the state of the text at one remove from the C-1 text type. I will argue that the prototype should be identified with the Prophet.

Chronology and the Role of the Prophet: The Argument from Stemmatics

The first argument about the identity of the prototype depends on the stemmata (p. 52-7). If one takes ʿUṯmān as a package (as opposed to a hybrid based on multiple Companion codices), then among the various stemmata,

(f) is the most plausible—this is the one in which the prototype P is dictated, yielding among others ʿUṯmān and C-1, which indeed exhibit the kinds of differences between themselves that would arise with dictation. Now, these codices go back to the time of the Companions. But the Companions were close to the Prophet temporally and, certainly in ʿUṯmān's case, personally. The common ancestor of their codices, the prototype P, would represent an even earlier stage of the text. But "an earlier stage of the text" when one is speaking of Companions does not leave much room for maneuver: P would most plausibly be the Prophet's own recension. As long as one acknowledges that the codices go back to the era of the Companions, the move to the Prophet as the prototype is not much of a leap.

The last paragraph proceeded on the basis of stemma (f). If one rejects (f), one shall be left with the less likely scenario (e) in which C-1 derived from ʿUṯmān, and that would push ʿUṯmān's textual tradition back in time to the period before C-1. But, given C-1's early date, in this scenario ʿUṯmān's version is most plausibly identified with the recension of the Prophet.

Two more scenarios: If one does not wish to acknowledge that there was a separate prototype, which is fine, then one must consider the remaining stemmata while removing P, but then it would still follow that ʿUṯmān was prior to C-1, which essentially makes ʿUṯmān itself the prototype. In terms of chronology, this scenario is no different than the ones just outlined. Finally, if one considers ʿUṯmān a hybrid, the stemma is (g). In this scenario, a number of Companion codices were compared to restore the wording of P. The early dates of all these Companion codices leads to the same conclusion as before.

Chronology and the Role of the Prophet: The Internal Evidence

Polarity analysis of internal evidence (p. 57-63) showed that either ʿUṯmān is a more reliable version of the ancestor it shares with C-1 (if one posits a P), or C-1 derives from ʿUṯmān. In light of the early dates of C-1 and ʿUṯmān, these scenarios again lead us back to the Prophet in a manner similar to the stemmata-based argument.

Learning about the Prophetic Prototype

The results of textual analysis can be compared with early Muslim reports. To the extent that they agree, there is independent corroboration. But where they are incompatible, the reports become questionable. Thus, textual analysis is a tool for evaluating early reports. It is particularly gratifying, however, that textual analysis can also shed light on matters on which literary sources

are silent or inconclusive. For example, historical reports, while suggestive, do not provide a strong basis for considering one text type as closer to the Prophet's own recension than another. Yet textual analysis enables one to probe this question. The study of the folios so far suggests that the text type of ʿUṯmān preserves the Prophetic prototype better than C-1.

This conclusion provides a broad characterization of the wording of the Prophetic prototype; but what can be inferred about its specifics? What one can say about P depends on which stemma one uses. If it is (g), then the work of textual analysis has already been performed for us, and it would be hard do better at recovering P than by simply reading ʿUṯmān. If it is (e), then ʿUṯmān is necessarily better than the other Companion codices. The rest of my discussion concerns the more challenging scenario of (f).

If the true stemma is (f), then the method is standard: commonalities in different branches (*i.e.* different text types) belong to their common ancestor; and in cases of disagreement, competing claims are weighed against one another: a majority vote among the branches establishes the original. The procedure and its outcome are most straightforward when all the text types agree. The general agreement of the branches on the order in which the verses are arranged into *sūra*s means that it was their common ancestor, *i.e.* the Prophetic prototype, that fixed the internal structure of the *sūra*s and the choice and order of the verses within them. In addition, the wording of the passages is generally the same in different codices, and common wording goes back to the Prophetic prototype. But what about the differences?

Going with majority vote, ʿUṯmān's text wins in the vast majority of cases, and the Prophetic prototype emerges as almost identical with the codex of ʿUṯmān. That is so because given any part of the text, either all three agree (which is the typical situation), or two agree against one, and in the latter case ʿUṯmān is usually in the majority. There are a number of exceptions, in which C-1 and Ibn Masʿūd agree against ʿUṯmān (see above, Table 4, p. 19).

But there is a problem with the procedure just described. The "one codex, one vote" motto may be egalitarian, but it is not just. If one source has been established as more reliable, then it should be given proportionately greater weight. Therefore, if at some point C-1 and IM agree against ʿUṯmān, it is not clear that they should override it. One must consider other possibilities, *e.g.* the Prophet having given revised versions of the prototype over time as Muslim tradition maintains (thus yielding two versions), cross-contamination, or, in certain cases, a concurrence of independent adaptations.[118]

[118] One may try to let reported variants of other Companion codices participate in the voting, but here one faces two challenges: determining which reports are reliable and, more daunting, ensuring that a reported variant is not a derivative of those that have already participated in the voting.

What about the verse divisions, *i.e.* the locations of the verse separators? Is it possible that the text was divided into verses for the first time after the Prophet? Or do the divisions go back to the Prophet? There is reason to answer the latter question in the affirmative: the verse divisions in the non-ʿUṯmānic C-1 agree with the ʿUṯmānic divisions, and therefore must have originated before either, *i.e.* with the Prophetic prototype. The application of X-Ray fluorescence imaging to Stanford '07 brought to light verse divisions not otherwise visible.[119] On that leaf, the locations of verse separators are identical to ʿUṯmānic locations (Iraqī system). Moreover, the location of C-1's 200th verse marker in *sūra* 2 is only two or three verses off compared to the ʿUṯmānic locations. This means C-1 and ʿUṯmānic verse divisions are usually the same over a stretch of some 200 verses. The match suggests a genetic link, *i.e.* common ancestor. It makes independent development unlikely since one can imagine alternative ways of separator placement: two independent attempts at dividing the verses would probably have led to different results.[120]

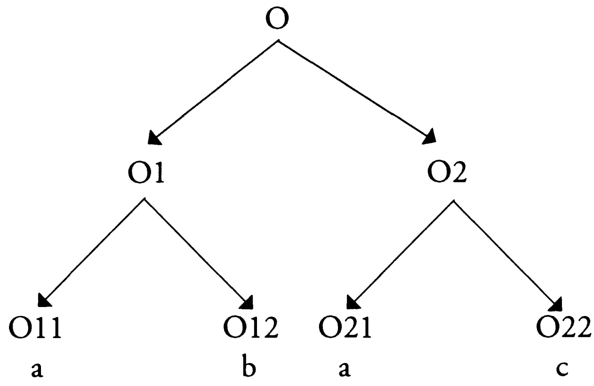
The fact that some ʿUṯmānic manuscripts have aberrant separator placements does not alter this conclusion. It is a principle of textual analysis that if different branches of a textual tradition include manuscripts with shared *distinctive* features, then the existence of manuscripts that do not share those features does not preclude assigning those shared features to the common ancestor.[121] The principle is illustrated in the diagram below. Symbols 'a', 'b', and 'c' represent a textual feature, say, a distinctive verse separator placement pattern. One ought to ascribe the feature 'a' to the common ancestor (O) even though it is not found in every manuscript. Because it is distinctive, its presence in the two independent branches is best described by genealogical descent from a common ancestor, and its absence in O12 and O22 by corruption. This conclusion would be suspended only if a good case were made for cross-contamination or independent convergent adaptations.

The above comments apply exclusively to the contents of the *sūra*s. But what about the order in which the *sūra*s are arranged relative to one another? Is it possible to infer the order in the Prophetic prototype? The evidence includes the *sūra* orders reported for Ibn Masʿūd and Ubayy b. Kaʿb. These reports must be taken seriously, in part because they are partially corroborated by C-1. In an attempt to ascertain the order of their common ancestor,

[119] When the attempt is made to study verse division practices in the entire extant portion of C-1, a caveat should be kept in mind: the fact that X-Ray fluorescence imaging brought into view verse separators that are not otherwise visible shows that in general if we do not see a verse separator in the lower text, that does not always mean the absence of one.

[120] Verse-separators on Stanford '07 that are "non-obvious" include those at Kor 2, 191, 198, 200, and 201. There are also spots without a verse separator at which one would have been imaginable.

[121] See, *e.g.*, the rules stated in Reynolds and Wilson, *Scribes & Scholars*, p. 211-2.



one may again conduct a vote. However, voting does not yield a definite result, since each codex has a different *sūra* order. The orderings do have commonalities, and these might go back to the Prophetic prototype; but the amount of disagreement is significant. The conclusion, then, is that the Prophet probably did not fix the order of the *sūra*s, except possibly in some cases. Such a conclusion is not new. Pre-modern scholars advanced a number of theories about whether the Prophet had fixed the order of the *sūra*s. The majority believed he had not.[122] Ibn ʿAṭiyya (d. 542/1148), however, suggested that the Prophet arranged some of the *sūra*s but not all.[123]

Implications for the nature of ʿU̱tmān's contribution

I make a key distinction between two types of evidence: (1) the collective, consistent, independent memory of the early communities on matters that we expect the communities to have known about, and (2) the anecdotes of individuals (especially on matters that the entire community could not be expected to have known about).[124] The former is reliable, while the latter requires evaluation. As discussed above, based (among other things) on the collective memory of communities, we know that ʿU̱tmān provided the cities' master codices. However, the exact details of how the ʿU̱tmānic text was

[122] Al-Suyūṭī, *al-Itqān*, p. 171.

[123] Ibn ʿAṭiyya, *al-Muḥarrar*, I, p. 50. My attention was drawn to this source by al-Aʿẓamī, *The History of the Qurʾānic Text*, p. 78. For Ibn ʿAṭiyya see, *e.g.*, Ḫayr al-Dīn al-Ziriklī, *al-Aʿlām: Qāmūs al-tarāǧim li-ašhar al-riǧāl wa-l-nisāʾ min al-ʿarab wa-l-mustaʿribīn wa-l-mustašriqīn*, 5th ed., Beirut, Dār al-ʿilm li-l-malāyīn, 1980, III, p. 282.

[124] Cf. Modarressi, "Early Debates", p. 13-4. For a discussion of why the collective memory can be considered reliable, see above p. 23-4.

put together was not something the entire community knew, and it was remembered in different ways. This is not surprising: the people of, say, Kūfa, as a whole could be expected to remember the name of the caliph who sent them their master codex, as the event would have been a very public affair, but relatively few would have had direct knowledge of the pre-history of that codex. As a result, there are conflicting trends in the traditions about the process that resulted in the standard text.

As Modarressi has noted, the very fact that ʿUṯmān standardized the Qurʾān suggests that there were different versions.[125] The burning question is how far-reaching their differences were. Did the codices arrange existing verses into *sūra*s in different ways, so that two versions of a *sūra* could have different verses, or have the same verses but order them differently? Or were the differences limited to the rather modest kind attested in C-1? More fundamentally, who arranged existing verses into *sūras*: was it ʿUṯmān's team or the Prophet? On this question there was no single collective memory, but rather divergent understandings handed down to us by individuals.

Two sets of reports of individuals support two very different positions. One group indicates that it was ʿUṯmān's team who strung existing verses together to form the *sūra*s, while another group ascribes this act to the Prophet.[126] Among reports that support ascription to the Prophet, one may include many of those describing the Companion codices: it would be meaningless to say that different Companions ordered the *sūra*s differently if the *sūra*s had not been formed already.[127] Furthermore, al-Aʿmaš's methodical collation of Ibn Masʿūd's verbal differences with ʿUṯmān and reports about other Companion codices presuppose that the codices were the same except for the specified differences at the level of words and phrases. The sequences of sentences and verses within the *sūra*s are the same in the known Companion codices, except for rare cases of transpositions of short adjacent verses.

Modern scholars, like pre-modern ones, have generally opted for one of these two positions. For example, Jeffery accepted the ascription to ʿUṯmān and therefore rejected the reports about the different *sūra* orderings of other Companions. On the other hand, Alford Welch and Hossein Modarressi

[125] Modarressi, "Early Debates", p. 14.

[126] Modarressi ("Early Debates") provides a brief survey of traditions suggesting that the *sūras* were fixed only after the Prophet's death (p. 8-13) before describing them as unhistorical. He calls them "extremely problematic" on p. 14 and goes on to propose a hypothesis explaining how they were generated. For another useful selection of the divergent traditions (though not a very convincing attempt to harmonize them), see al-Aʿẓamī, *The History of the Qurʾānic Text*, p. 71-101.

[127] A.T. Welch, R. Paret, J.D. Pearson, "al-Ḳurʾān", *EI*².

favor the idea that the *sūra*s were put in their final forms during the Prophet's lifetime. The study of C-1 has confirmed the latter position.

Implications for the Historicity of Companion Codices

Manuscript evidence now corroborates pre-modern reports about the existence of Companion codices, their having different *sūra* orderings, and, to an extent, the nature of their verbal differences. Conclusively refuted is John Burton's theory that all such reports were post-ʿUṯmānic fictions aimed at "countering, ellucidating, or even evading the ʿUṯmān text."[128] But once one acknowledges the traditional notion of Companion codices, it remains to explain their differences. Such differences would have arisen as the Prophet recited the *sūra*s and different scribes wrote them down.

Given that believers were urged to recite the Qur'ān as much as feasible (Kor 73, 20), the question arises whether the Prophet knew of some of these differences, and, if he did, what his reaction was. Ibn Masʿūd may or may not have spoken the words ascribed to him in the following report, but there is no doubt that they embody one of the earliest theories explaining the differences among the Companion codices:

> Abū Usāma—Zuhayr—al-Walīd b. Qays—ʿUṯmān b. Ḥassān al-ʿĀmirī—Filfila l-Ǧuʿfī: he [*i.e.* this last] said: I was among those who went to ʿAbd Allāh [Ibn Masʿūd], fearful over [the issue of] the codices. We came into his presence and a man in the group said, "We have not come to you as guests; we came here upon being frightened by this news [of the standardization]." He answered, "The Qur'ān was sent down to your prophet through seven [heavenly] gates according to seven modes (*aḥruf*, *ḥurūf*). Previous to you [*i.e.* before your generation], the Book was sent down through a single gate according to a single mode (*ḥarf*).[129]

The report presupposes that the Prophet approved of Ibn Masʿūd's version of the Qur'ān alongside that promulgated by ʿUṯmān. It is not inconceivable that different scribes read different versions back to the Prophet, and were

[128] John Burton, *The Collection of the Qur'an*, Cambridge, Cambridge University Press, 1977, p. 228. Burton assumed that if one can imagine a theological motive for creating a report, the report cannot be historically true. The problem is that often with a little imagination one can concoct theological motives even for accurate reports.

[129] Ibn Abī Dāwūd, *al-Maṣāḥif*, p. 18. Edward Lane, when defining *ḥarf* in his *Lexicon*, cites this expression: *fulān yaqraʾu bi-ḥarf Ibn Masʿūd*. From the report quoted by Ibn Abī Dāwūd and other evidence, it appears that the "modes" (*aḥruf*) in which the Qur'ān is said to have been revealed originally encompassed the Companion codices. This also happens to have been the position of most premodern and early-modern scholars. See above, p. 4 and footnote 4.

met with his tacit approval when they did so.[130] Yet, it must be stressed that at present there is no definitive evidence for or against this. In any case, if the Prophet did tacitly approve more than one version, that would not necessarily mean that all versions represented equally precisely the recitations as they left his mouth.

Summary

Muḥammad dictated the revelations, and scribes wrote them down. This gave rise to a number of Companion codices. As the Prophet had not fully determined the order of the *sūras* relative to one another, these codices had different *sūra* orders. However, he had fixed the contents of the *sūras*, including the distribution of verses within them and even the verse divisions. On these elements, and especially where the actual text is concerned, the codices showed great agreement.

Yet, the aural dimension of the Prophet's dictation at times generated changes, giving rise to occasional verbal differences. Many of the differences among the Companion codices point to semi-orality, and they go back to the Companions' transcription of a Qur'ān recited by the Prophet. If the scribes recited the text back to the Prophet—and we do not know whether this happened—one wonders if the Prophet tacitly endorsed some of these differences, relatively small as they generally seem to be. If so, that would not negate the fact that one version better represented what the Prophet himself actually recited; but which one?

The caliph ʿUṯmān established the standard version, an undertaking that, according to the literary sources, involved a committee, and, above all, a scribe of the Prophet named Zayd b. Ṯābit.[131] If ʿUṯmān formed a committee to deal with this potentially explosive issue, then that was a politically astute move, making it easier to gain the acceptance of a large part of the community, and helping deflect criticism from the caliph himself.[132] And if it is true

[130] This scenario fits a perhaps *improbable* reading of Kor 25, 5. The verse runs: *wa-qālū asāṭīru l-awwalīna ktatabahā fa-hiya tumlā ʿalayhi bukratan wa-aṣīlā*. It says of Muḥammad's opponents in Mecca that, "They say, '[that the Qur'ān is] fairy-tales of the ancients that he [Muḥammad] has caused to be written down, so that they are recited (or dictated) to him at the dawn and in the evening.'" This translation closely follows those of Arberry, Asad, Sher Ali, and Abdalhaqq and Aisha Bewley. The improbable reading would be that the Prophet would have followers write the Qur'ān, by dictation, and it would be recited back to him (*tumlā ʿalayhi*). The perhaps more probable reading is that enemies charged that informants recited the tales to the Prophet day and night. Either way, the implication is that the Qur'ān was written.

[131] Rāmyār, *Tārīḫ-i Qur'ān*, p. 416-35; A.T. Welch, *et al.*, "al-Ḳur'ān".

[132] Michael Cook, "Concluding Remarks", Colloquium on the Evidence for the Early

that the outcome was a hybrid codex, then that had the added political advantage that Companions would not be chagrined to learn that a fellow Companion's codex was preferred over theirs.[133]

In any case, textual criticism suggests that the standard version is the most faithful representation, among the known codices, of the Qur'ān as recited by the Prophet. This appears, at first, as a curious coincidence; but on second thought it is not surprising: if anybody had the resources to ensure that a reliable version be chosen, it would have been the caliph; and if anybody had more to lose by botching up the task, again that would have been ʿUṯmān, whose political legitimacy and efficacy as caliph depended completely on the good will of fellow distinguished associates of the Prophet. The remarkably few and minor skeletal-morphemic differences among the codices ʿUṯmān sent to the cities is another indication of the care that was put into the process of standardization.[134]

Not everybody considered the standard version as the only legitimate one. Some, espousing a sort of codical pluralism, continued to consider other Companion codices as legitimate. The codex of Ibn Masʿūd, in particular, continued to have supporters in the first two centuries AH. Nonetheless, upon its dissemination the standard version quickly became predominant everywhere. Given the vast expanse and decentralized nature of the empire, the center's intervention had achieved a remarkable degree of success. Several years after Medina's act of standardization, however, a new era was ushered in. The murder of ʿUṯmān in AD 656 occasioned what came to be known as the "First Civil War" in the historical memory of Muslims. It polarized and fractured the community irrevocably. Had empire-wide standardization been attempted at any moment after this point, it would have been a hopeless undertaking.

Post-Script: Directions for Future Research

It is a curious and telling fact of the field of Qur'ānic studies that few of its achievements in the West have been celebrated more than a project that never gathered much steam to begin with, namely the plan by Bergsträsser and his

---

History of the Qur'ān, Stanford University, July 31, 2009. Rāmyār, too, points out that Ibn Masʿūd's attacks were directed against Zayd (*Tārīḫ-i Qur'ān*, p. 418).

[133] Rāmyār, *Tārīḫ-i Qur'ān*, p. 422.

[134] The variants typically affect one or two characters without changing the meaning, and they number thirty-six. As Cook notes, "This is not a large number for a text as long as the Koran; the proportion of words affected is less than one in two thousand. The number of distinct readings at any given point never exceeds two, each regional codex having one or the other" (Cook, "The Stemma of the Regional Codices of the Koran", p. 90-1).

collaborators to prepare a critical text of the scripture.[135] And it is perhaps in keeping with this unreal dimension that now we learn that the oft-repeated reason for its termination was a deliberate lie: the tale was that Allied bombing during WWII destroyed a collection of photographs of early *qirāʾāt* works and photographs of Qurʾāns in the Kūfī script. Yet in the midst of the elegies in the past decades that put the blame on bombs, few seemed to notice that early Qurʾān manuscripts that *have* been available and accessible in Europe and the Muslim world were not being studied.[136] In fact, only in the last few years have some of them been subjected to textual analysis by scholars such as Yasin Dutton. Nor was it apparent to all that if the study of *qirāʾāt* moved at a glacial pace, this had nothing to do with lack of materials.

It is obvious that the study of the early history of the Qurʾān requires a dual track involving (1) the *qirāʾāt* and early traditions, and (2) manuscripts. This two-pronged approach was, for example, what Arthur Jeffery saw over eighty years ago as the basis of his collaboration in the above-mentioned project. A preliminary step would have been to collect the "textual variants gathered from the Commentaries, Lexica, *Qirāʾāt* books and such sources",[137] a task that would be "superhuman" if done in a comprehensive manner.[138]

The dream that one day all the *qirāʾāt* variants would be collected has been fulfilled, although the fulfillment has not received, in the West, anything like the recognition accorded to Bergsträsser's and Jeffery's dream. Many *qirāʾāt* works have been published and some have been analyzed in the Muslim world. There are now also encyclopedias: the *Muʿǧam al-qirāʾāt al-qurʾāniyya* of Aḥmad Muḫtār ʿUmar and ʿAbd al-ʿĀl Sālim Mukarram has gone through two editions in Kuwait and at least one printing in Iran.[139] In addition, ʿAbd al-Laṭīf Muḥammad Ḫaṭīb's comprehensive *Muʿǧam al-qirāʾāt* was published in Syria. It is an erudite work that canvasses works of exegesis, grammar, and *qirāʾāt*. These efforts have mapped the literature and paved the way for analytical studies. Building on them is a desideratum of the field of Qurʾānic studies.

As for the study of manuscripts, it may appear to some that what the field needs most is a database of images of all Qurʾān manuscripts. That would undoubtedly be an invaluable resource, but let us be clear about what it could

---

[135] For a summary of the critical text project, see Jeffery, *Materials*, p. *vii*.

[136] It is the lack of systematic textual studies of manuscripts to which I am referring. There have been useful paleographic studies by Abbott, Munajjid, Déroche, Whelan, and others. See the sources cited in the works listed in footnote 50.

[137] Jeffery, *Materials*, p. *vii*.

[138] Jeffery, as quoted in Rabb, "Non-Canonical Readings", p. 87.

[139] Aḥmad Muḫtār ʿUmar, ʿAbd al-ʿĀl Sālim Mukarram, *Muʿǧam al-qirāʾāt al-qurʾāniyya: maʿa muqaddima fī l-qirāʾāt wa-ašhar al-qurrāʾ*, 2nd edition, Kuwait, Ḏāt al-Salāsil, 1988.

or could not help achieve. Obviously, non-ʿUṯmānic manuscripts would tell us much about the early history of the Qurʾān, but so far there is only one known manuscript in this category, namely Ṣanʿāʾ 1. When one speaks of collecting all manuscripts, one means essentially the innumerable ʿUṯmānic materials. The main aim of this gargantuan task would be to reconstruct their archetype, the original codex of ʿUṯmān. But as explained above (p. 26), the work of Michael Cook and others has shown that we already know the skeletal-morphemic text of the original codices of ʿUṯmān, meaning that the uncertainty of the skeletal text primarily concerns spelling. A list of all the ways in which the words in ʿUṯmānic manuscripts differ from that known text would be nothing more than a list of later accretions, useful at best for tracing the history of manuscript production after AD 656. New information might clarify certain spelling practices, some disputed undotted consonants and disputed verse divisions in the original codices sent out by ʿUṯmān; but while any new tidbit would be exciting for the Qurʾān specialist, rarely will such information make a major difference in our understanding of the history of the Qurʾān.

The amount of work yet to be done is great, and the main paths of embarking on the tasks are clear. It is now equally clear that recent works in the genre of historical fiction are of no help. By "historical fiction" I am referring to the work of authors who, contentedly ensconced next to the mountain of material in the premodern Muslim primary and secondary literature bearing on Islamic origins, say that there are no heights to scale, nothing to learn from the literature, and who speak of the paucity of evidence. Liberated from the requirement to analyze the literature critically, they can dream up imaginative historical narratives rooted in meager cherry-picked or irrelevant evidence, or in some cases no evidence at all. They write off the mountain as the illusory product of religious dogma or of empire-wide conspiracies or mass amnesia or deception, not realizing that literary sources need not always be taken at face value to prove a point; or they simply pass over the mass of the evidence in silence. A pioneering early example of such historical fiction was *Hagarism*, written by Patricia Crone and Michael Cook. While few specialists have accepted its narrative, the book has nevertheless profoundly shaped the outlook of scholars. It has given rise to a class of students and educators who will tell you not only that we *do not* know anything about Islamic origins, but also that we *cannot* learn anything about it from the literary sources. All this would be good and well if the mountain of evidence had been studied critically before being dismissed as a mole hill; but the modern critical reevaluation of the literary evidence has barely begun. And, significantly, any number of results have already demonstrated that if only one

takes the trouble to do the work, positive results *are* forthcoming, and that the landscape of the literary evidence, far from being one of randomly-scattered debris, in fact often coheres in remarkable ways. A good example of such findings would be some of Michael Cook's own fruitful recent studies in the literary sources in two essays of his already discussed here. It is not his confirmation of some elements of the traditional account of the standard Qur'ān that I wish to highlight here, noteworthy as it may be, but rather his demonstration that we *can* learn from the study of the literary sources.

Appendix 1. Lower Text: (a) Minor Elements and (b) Phonetic Conservation of Major Elements

Items placed in parentheses (x) are only partly visible, but enough is visible to have good reason for the readings given here.

Items placed in double parentheses ((x)) are not visible, and the readings offered here are usually largely speculative.

(a) *Minor Elements*

For the definition of "minor", see above, p. 44.

Verse	ʿUṯmān	C-1	Type of Variant
2, 200	ما له	ما لهم	minor (person)

Note: the C-1 version is problematic. The personal pronoun refers to *man*, and it should be singular.

Verse	ʿUṯmān	C-1	Type of Variant
2, 210	ينظرون إلا أن يأتيهم	تنظرون إلا أن يأتيكم	minor (person)
2, 220	فإخوانكم	فإخوانهم	minor (person)

Note: the C-1 version makes little sense, while the version of ʿUṯmān reads naturally.

Verse	ʿUṯmān	C-1	Type of Variant
5, 45	له	(لكم)	minor (person)
5, 51	بعضهم	بعضكم	minor (person)
63, 1	نشهد	يشهدو(ا)	minor (person)
63, 4	فاحذرهم	((فاحذروهم))	minor (person)

(*cont.*)

Verse	ʿUṯmān	C-1	Type of Variant
63, 10	...فيقول رب أخرتني...فأصدق وأكن	فيقول ربنا . . . أخرنا فأصدق (وأكن)	minor (person)
62, 3	منهم...بهم	(منكم...بهم)	minor (person)
2, 191	يقتلوكم فيه	يقتلوكم	minor omission (*fīhi*)
2, 196	فإن أحصرتم	إن أحصرتم	minor omission (*fa*)
2, 196	ففدية	فدية	minor omission (*fa*) Apparent error of the hand
2, 198	من قبله	قبله	minor omission (*min*)
2, 217	سبيل الله	(سبيله)	minor omission (*allāh*)
5, 42	أعرض عنهم	أعرض	minor omission (*ʿanhum*)
5, 43	إنا أنزلنا	وأنزلنا	minor omission (*innā*)
5, 49	وأحذرهم أن	(وأحذر أن)	minor omission (*hum*)
15, 59	انا لمنجوهم	انا منجوهم	minor omission (*la*)
63, 1	لرسوله	((رسوله))	minor omission (*la*)
62, 11	من اللهو ومن التجرة	من اللهو والتجرة	minor omission (*wa-min*)
89, 25	فيومئذ	((يومئذ))	minor omission (*fa*)
2, 194	بمثل ما اعتدى عليكم	بمثل ما اعتدى عليكم به	minor addition (*bihi*)
2, 217	قتال فيه	و(عن) قتل فيه	minor addition (*wa-ʿan*)
2, 197	فلا رفث	فلا رفث ((فيهن))	minor addition (*fīhinna*)

C-1's version is a candidate for the assimilation of a nearby term.

(*cont.*)

Verse	ʿUṯmān	C-1	Type of Variant
2, 204	في الدنيا	في ((هذه)) الدنيا	minor addition (*hāḏihi*)

C-1's version is a candidate for the assimilation of parallels at Kor 3, 117; 7, 156; 11, 60; 16, 30; 20, 72; 28, 42; 29, 64; 39, 10; and 40, 39.

2, 194	فمن	ومن	minor substitution (*fa* vs. *wa*)
2, 196	بالعمرة	بعمرته	minor substitution (possessive vs. definite)
2., 98	فضلا	الفضل	minor substitution (definite vs. indefinite)
2, 200	فإذا	وإذا	minor substitution (*fa* vs. *wa*)
5, 42	فلن يضروک	لا يضروک	minor substitution (*lan* vs. *lā*)

For a similar substitution, but in the opposite direction, see below, Kor 62, 7.

5, 43	وما اولئك	فما اولئك	minor substitution (*wa* vs. *fa*)
5, 44	للذين	والذين	minor substitution (*li* vs. *wa*)
5, 54	فسوف	فسـ	minor substitution (synonyms *sawfa* vs. *sa*)
63, 8	ولرسوله وللمؤمنين	(وللرسول والمؤمنين)	minor substitution (possessive vs. definite)
62, 7	ولا يتمنونه	ولن يتمنوه	minor substitution (*lā* vs. *lan*)

For a similar substitution, but in the opposite direction, see above, Kor 5, 42.

(*cont.*)

(b) *Phonetic Conservation of Major Elements*

For the definition of "major", see above, p. 45.

Verse	ʿUṯmān	C-1	Type of Variant
2, 196	استيسر	تيسر	phonetic conservation (same-root synonym)
2, 196	فما استيسر	فما تيسر	phonetic conservation (same-root synonym)
2, 219	الأيت	((الأية))	phonetic conservation (singular vs. plural)

The plural form (in ʿUṯmān) reads more naturally.

Elsewhere in the Qurʾān, the phrase *yubayyinu llāhu* is always followed by the plural form *āyāt*. See Kor 2, 187; 2, 219; 2, 242; 2, 266; 3, 103; 5, 89; 24, 18; 24, 58; 24, 59; 24, 61.

2, 222	فإذا تطهرن	فإذا (طهرن)	phonetic conservation (same-root verb)
5, 41	ولهم في الاخرة	و(في) (الاخر)ة لهم	phonetic conservation (word order: transposition)
5, 42	سمعون...أكلون	أكلون... (يسمعون)	phonetic conservation (same-root alternative)
5, 43	وكيف يحكمونك وعندهم التورية فيها حكم الله	وكيف يحكمونك والتورية فيها حكم الله (عندهم)	phonetic conservation (word order: transposition)
5, 43	هدى ونور	نور و(هدى)	phonetic conservation (word order: transposition)

(*cont.*)

Verse	ʿUṯmān	C-1	Type of Variant
5, 43	النبيون	الأنبياء	phonetic conservation (equivalent plural forms)

Both forms appear in the standard Qurʾān.

5, 45	من تصدق به	(فإن) تصدقوا بها	phonetic conservation (syntactical structure plus minor elements, person and gender)

In C-1, due to illegibility, I cannot tell if what follows is *fa-huwa* or *fa-hiya*. ʿUṯmān's term *bihi* seems more appropriate than C-1's term *bihā*.

For another instance of C-1's propensity for constructing conditionals with *in* instead of *man*, see below, Appendix II (c), Kor 2, 196. Both structures are well-attested in the Qurʾān.

5, 47	أنزل الله فيه	أنزل (عليهم) فيه	phonetic conservation (active vs. passive, plus different minor elements)

The passive form *unzila* followed by the preposition *ʿalā* occurs eleven times (Kor 2, 91; 3, 84; 6, 8; 6, 157; 10, 20; 11, 12; 13, 7; 13, 27; 25, 21; 29, 50; 38, 8).

The active form *anzala* occurs some sixty-six times, including twenty-nine times followed immediately by *llāh*.

5, 48	لكل جعلنا منكم شرعة	لكل ((منكم)) جعلنا (شرعة)	phonetic conservation (word order: transposition) or minor omission
5, 50	من الله حكما	(حكما) من الله	phonetic conservation (word order: transposition)

(*cont.*)

Verse	ʿUṯmān	C-1	Type of Variant
5, 52	فيصبحوا	فيصبحو(ن)	phonetic conservation (verb in different aspect)
15, 65	تؤمرون	يامرون	phonetic conservation (active vs. passive, plus person)
63, 1	والله يشهد	ويشهد (الله)	phonetic conservation (word order: transposition)
62, 7	يتمنونه	يتمنوه	phonetic conservation (*yatamannawnahu* vs. *yatamannawhu*)

As for ʿUṯmān's version, the word *yatamannawnahu* is found nowhere else in the Qurʾān, although in Kor 3, 143 we have *kuntum tamannawna l-mawta.*

As for C-1, the entire verse here is found also in Kor 2, 195 with C-1's wording, *yatamannawhu.*

Appendix 2. Lower Text, Major Elements: (a) Omissions, (b) Additions, (c) Substitutions, (d) Scribal errors, (e) Illegible variants

Items placed in parentheses (x) are only partly visible, but enough is visible to have good reason for the readings given here.

Items placed in double parentheses ((x)) are not visible, and the readings offered here are usually largely speculative.

(a) *"Omissions" in C-1 (i.e. pluses of ʿUṯmān)*

Verse	ʿUṯmān	C-1	Type of Variant (C-1 relative to ʿUṯmān)
2, 196	ولا تحلقوا رؤسكم	ولا تحلقوا	omission (*ruʾūsakum*)

(*cont.*)

Verse	ʿUṯmān	C-1	Type of Variant (C-1 relative to ʿUṯmān)

The version in ʿUṯmān has something of a parallel elsewhere: the word *ruʾūsakum* and the verb *ḥalaqa* occur only in one other spot in the Qurʾān, at Kor 48.27, in *muḥalliqīna ruʾūsakum*.

2, 196	من صيام أو صدقة أو نسك	من صيم أو نسك	omission in a list (*aw ṣadaqatin*)

ʿUṯmān has *min ṣiyāmin aw ṣadaqatin aw nusukin*, whereas C-1 gives *min ṣiyāmin aw nusukin*.

Neither version has parallels elsewhere in the Qurʾān.

The omitted word in C-1, *ṣadaqa*, occurs six times in ʿUṯmān (Kor 2, 196; 2, 263; 4, 114; 9, 103; 12, 26; 58, 12). Its plural, *ṣadaqāt*, occurs eight times (2, 264; 2, 271; 2, 276; 9, 58; 9, 60; 9, 79; 9, 104; 58, 13).

2, 201	في الدنيا حسنة وفي الآخرة حسنة وقنا	في الدنيا والآخرة وقنا	omission (*ḥasana… fī… ḥasana*)

The closest parallel is Kor 7, 156: *wa-ktub lanā fī hāḏihi l-dunyā ḥasanatan wa-fī l-āḫirati*; cf. 16, 30; 16, 41 and 16, 122. The phrase "this world and the next" (*l-dunyā wa-l-āḫira*) occurs fifteen other times in the Qurʾān. The word *ḥasana* occurs twenty-six additional times in the Qurʾān.

Neither the version in ʿUṯmān nor that in C-1 can be considered unexpected, and both make sense in context.

2, 213	بغيا بينهم		omission (*baġyan baynahum*)

ʿUṯmān's phrase *baġyan baynahum* appears in three other spots in similar contexts: Kor 3, 19; 42, 14; 45, 17. The term *baġyan* is also used alone in 2, 90 and 10, 90.

2, 213	الذين خلوا من قبلكم	الذين من قبلكم	omission (*ḫalaw*)

(*cont.*)

Verse	ʿUṯmān	C-1	Type of Variant (C-1 relative to ʿUṯmān)

The text of ʿUṯmān here has several parallels: Kor 24, 34 (*allaḏīna ḫalaw min qablikum*); 10, 102 (*allaḏīna ḫalaw min qablihim*), and 33, 38 and 33, 62 (*fī llaḏīna ḫalaw min qablu*).

However, the C-1 version is much more richly attested. We have the phrase *allaḏīna min qablikum* in Kor 2, 183; 4, 26; 9, 69; 14, 9; *allaḏīna min qablihim* in Kor 2, 118; 6, 148; 9, 70; 10, 12; 12, 109; 13, 42; 16, 26; 16, 33; 16, 35; 24, 55; 24, 59; 29, 3; 30, 9; 34, 45; 35, 25; 35, 44; 39, 25; 39, 50; 40, 82; 47, 10; 51, 52; 58, 5; 59, 2; 67, 18; *allaḏīna min qablinā* in 2, 286; *allaḏīna min qablu* in 30, 42; and *allaḏīna min qablika* in 39, 65 and 42, 3.

Verse	ʿUṯmān	C-1	Type of Variant (C-1 relative to ʿUṯmān)
2, 217	وكفر به		omission (*wa-kufrun bihi*)

ʿUṯmān's *kufrun bihi* is the only instance of this phrase in the entire Qurʾān. So, one cannot consider it as expected in context.

Verse	ʿUṯmān	C-1	Type of Variant (C-1 relative to ʿUṯmān)
2, 217	يردوكم عن دينكم	يردوكم	omission (*ʿan dīnikum*)

ʿUṯmān has the specifying phrase *ʿan dīnikum* ("from your religion"), but C-1 does not. Elsewhere in the Qurʾān, the verb *yaruddūkum* ("make you turn") and its cognates may occur with or without the specifying phrase "from your/their/his religion". The specifying phrase is included, for example, in *both* C-1 and ʿUṯmān in Kor 5.54 as *yartadda minkum ʿan dīnikum* (cf. Kor 2, 217). However, when the verb occurs without the specifying phrase, as in 2, 109, 3, 100 and 3, 149, the context makes it clear what the person is returning from. In other words, at first sight C-1 appears anomalous at the present point for not specifying what one would turn away from. However, the presence in C-1, later in the verse, of *yartadid minkum ʿan dīnihi* as established by the partly visible *nūn* of *dīnihi*, can be said to remove ambiguity.

Verse	ʿUṯmān	C-1	Type of Variant (C-1 relative to ʿUṯmān)
2, 218	ان الذين امنوا والذين هاجروا وجهدوا	ان الذين امنوا وهجروا	omission in a list (*wa-llaḏīna… wa-ǧāhadū*)

(*cont.*)

Verse	ʿUṯmān	C-1	Type of Variant (C-1 relative to ʿUṯmān)

ʿUṯmān has *inna llaḏīna āmanū wa-llaḏīna hāǧarū wa-ǧāhadū*, whereas C-1 gives *inna llaḏīna āmanū wa-hāǧarū*. Essentially, ʿUṯmān has an item in the list that C-1 lacks, namely *ǧāhadū*. The other elements, found in both texts, are *āmanū* and *hāǧarū*.

Elsewhere in the standard Qurʾān, *hāǧarū* never appears with *āmanū* unless *ǧāhadū* is also present. Therefore, their juxtaposition of the pair in C-1 cannot be labeled expected.

As for ʿUṯmān's version, we find the three verbs of its list are juxtaposed also at three other spots, in Kor 8, 72; 8, 74 and 9, 20, as *llaḏīna āmanū wa-hāǧarū wa-ǧāhadū* (cf. Kor 8, 75). Therefore, the presence of *ǧāhadū* at the present spot would not be unexpected.

2, 221	المغفرة بإذنه	المغفرة	omission (*bi-iḏnihi*)

ʿUṯmān's *bi-iḏnihi* ("by his leave") occurs also in eight other spots in the Qurʾān (2, 213; 2, 255; 3, 152; 5, 16; 11, 105; 22, 65; 33, 46; 42, 51), and *bi-iḏni* ("by the leave of") occurs twenty-three times followed by *llāh* or *rabb*. In none of these cases is the term mentioned in conjunction with forgiveness (*maġfira*). Nonetheless, the verse at hand finds a parallel in Kor 33, 46 (*wa-dāʿiyan ilā llāhi bi-iḏnihi*). Due to this parallel, we may say that the version of ʿUṯmān would not be entirely unexpected.

2, 222	فاعتزلوا النساء في المحيض ولا تقربوهن حتى يطهرن	(فلا تقربوا) النساء في (محيضهن) حتى يتطهرن	omission (*fa-ʿtazilū*) in addition to word order & minor elements.

The key difference here is the presence of the imperative *fa-ʿtazilū* in ʿUṯmān and its absence in C-1. (The other differences are reducible to phonetic conservation and minor elements.) The verb *iʿtazala* occurs at six other points in the Qurʾān, including only once as an imperative, but never in a marital or sexual context. There are absolutely no parallels for the verse at hand. Therefore, the verb *fa-ʿtazilū* is not something that a scribe would have anticipated here. ʿUṯmān's version is not a candidate for auto-contamination. See Kor 4, 90; 4, 91; 18, 16; 19, 48; 19, 49; 44, 21.

(*cont.*)

Verse	ʿUṯmān	C-1	Type of Variant (C-1 relative to ʿUṯmān)
5, 42	فإن جاءوك فاحكم بينهم	فاحكم بينهم	omission (*fa-in ǧāʾūka*)

The phrase in ʿUṯmān is not attested elsewhere. The word *ǧāʾūka* appears four other times in the Qurʾān (Kor 4, 62; 4, 64; 6, 25; 58, 8), twice preceded by *iḏā*, but nowhere by *in*. The word *ǧāʾaka* occurs eleven times in the standard Qurʾān, preceded three times by *iḏā*. See 2, 120; 2, 145; 3, 61; 5, 48; 6, 34; 6, 54; 10, 94; 13, 37; 60, 12; 63, 1; 80, 8. The word جاءوك also occurs in C-1 at Kor 63, 4, where it harks back to the *ǧāʾaka* at the beginning of that *sūra* (see below).

Verse	ʿUṯmān	C-1	Type of Variant
5, 44	فلا تخشوا الناس	فلا تخشوهم	omission (*l-nās*)

ʿUṯmān's phrase *fa-lā taḫšawu l-nāsa* occurs nowhere else in the Qurʾān, although it has parallels in Kor 4, 77 (*farīqun minhum yaḫšawna l-nāsa ka-ḫašyati llāhi*) and 33, 37 (*wa-taḫšā l-nāsa wa-llāhu aḥaqqu an taḫšāhu*). The word *al-nās* is a common word in the Qurʾān, occurring 247 times.

On the other hand, C-1's phrase *fa-lā taḫšawhum* is attested in Kor 2, 150 and 5, 3.

Verse	ʿUṯmān	C-1	Type of Variant
5, 49	فاعلم أنما	(فإنما)	omission (*fa-ʿlam*)

ʿUṯmān has *fa-in tawallaw fa-ʿlam annamā yurīdu llāhu an yuṣībahum bi-baʿḍi ḏunūbihim*. ʿUṯmān's juxtaposition of the verbs *tawallaw* (in a conditional clause) and *fa-ʿlam* also occurs in Kor 5, 92 (*fa-in tawallaytum fa-ʿlamū*), Kor 8, 40 (*wa-in tawallaw fa-ʿlamū*), Kor 9, 3 (*wa-in tawallaytum fa-ʿlamū*)

On the other hand, the version of C-1 has even more parallels: Kor 2, 137 (*wa-in tawallaw fa-innamā*), 3, 20 (same), 16, 82 (same, with *fa*), 24, 54 (same, with *fa*), 64, 12 (*fa-in tawallaytum fa-innamā*).

Verse	ʿUṯmān	C-1	Type of Variant
63, 1	المنفقون قالوا	المنفقون	omission (*qālū*)

The word *qālū* is ubiquitous in the Qurʾān.

(*cont.*)
(b) *"Additions" in C-1 (i.e. pluses of C-1).*

Verse	ʿUṯmān	C-1	Type of Variant
2, 193	ويكون الدين لله	ويكون الدين كله لله	addition (*kulluhu*)

The entire sentence in C-1 occurs elsewhere in ʿUṯmān, in Kor 8, 39. So, the word *kulluhu* would have been expected here for somebody familiar with the language of the Qurʾān.

ʿUṯmān's version, on the other hand, appears nowhere else.

2, 200	اتنا في الدنيا	اتنا في (هذه الحيوة) الدنيا	addition (*hādhihi l-ḥayāt*)

C-1's version is a candidate for the assimilation of parallels at Kor 3, 117; 7, 156; 11, 60; 16, 30; 20, 72; 28, 42; 29, 64; 39, 10; and 40, 39.

5, 44	للذين هادوا والربنيون والاحبار بما استحفظوا من كتب الله وكانوا عليه شهداء	والذين هدوا يحكمون بما نزل الله فيها ((و)) يحكم به اا((ربنيـ)) ون و الا((حبر)) بما استحفظوا من كتب الله ((و)) كانوا عليه شهداء	addition (*yaḥkumūna bi-mā nazzala llāhu fīhā* and *yaḥkumu bihi*)

C-1's *yaḥkumūna bi-mā nazzala* (read *anzala*?) *llāhu fīhā* is closely paralleled by another phrase in the same verse: *man lam yaḥkum bi-mā anzala llāhu*, a phrase that repeats two more times in the next two verses. Then in the verse after those (5, 48) we have *fa-ḥkum baynahum fī-mā anzala llāhu*. A similar phrase appears in the next verse as well (5, 49). Therefore, the additional terms in C-1 mirror language that repeats five times in close proximity to the spot in question, and therefore would not be "unexpected". It is a candidate for assimilation of nearby phrases.

5, 45	وكتبنا عليهم	وكتبنا على بني اسرائل	addition (*ʿalā banī Isrāʾīla*)

The phrase *katabnā ʿalā banī Isrāʾīla* occurs in Kor 5, 32, *i.e.* shortly before the verse at hand. The proximity of this occurrence makes it easy to see how the addition in C-1 could be a false memory harking back to the earlier use.

(*cont.*)

Verse	ʿUṯmān	C-1	Type of Variant
63, 3	ذلك بأنهم ءامنوا ثم كفروا	ذلك بأنهم ءامنوا ثم كفروا (ثم)ا(ز)ددوا (كفر)ا	addition (*ṯumma zdādū kufran*)

The version in C-1 is a restatement of Kor 4, 137 (*ṯumma āmanū ṯumma kafarū ṯumma zdādū kufran*). Cf. 3, 90 and 3, 178.

63, 3	ذلك بأنهم... فطبع على قلوبهم فهم لا يفقهون	ذلك بأنهم... فطبع على قلوبهم ذلك بأنهم قوم لا يفقهون	addition (*ḏālika bi-annahum qawmun*)

C-1's phrase *ḏālika bi-annahum* is simply a repetition of the same term earlier in the verse. The construct *qawmun lā* occurs seven times (Kor 5, 58; 8, 65; 9, 6; 9, 127; 43, 88; 59, 13; 59, 14), including three times as *qawmun lā yafqahūna* (Kor 8, 65; 9, 127; 59, 13).

On the other hand, ʿUṯmān's phrase *fa-hum lā yafqahūna* is also attested; see Kor 9, 87. There are also other instances where *lā yafqahūna* is used without a preceding *qawm*, *viz.* Kor 7, 179; 48, 15; 63, 7.

63, 7	ينفضوا	ينفضوا من حوله	addition (*min ḥawlihi*)

The version of C-1 has a parallel in Kor 3, 159 (*la-nfaḍḍū min ḥawlika*).

63, 8	ولله العزة	ولله العزة جميعا	addition (*ǧamīʿan*).

C-1's version has the word *ǧamīʿan*, which occurs in the Qurʾān forty-nine times. For close parallels of C-1's expression, see 35, 10 (*fa-li-llāhi l-ʿizzatu ǧamīʿan*); 4, 139 (*fa-inna l-ʿizzata li-llāhi ǧamīʿan*); 10, 65 (*inna l-ʿizzata li-llāhi ǧamīʿan*); 13, 31 (*li-llāhi l-amru ǧamīʿan*); 39, 44 (*li-llāhi l-šafāʿatu ǧamīʿan*); 2, 165 (*anna l-quwwata li-llāhi ǧamīʿan*).

As for ʿUṯmān's phrase *wa-li-llāhi l-ʿizzatu*, nowhere else does it occur in this form; elsewhere it appears only with *ǧamīʿan* as in C-1's version here.

63, 11	إذا	(لغد)إن	addition (*li-ġadin*)

(*cont.*)

Verse	ʿUṯmān	C-1	Type of Variant

The C-1 word *li-ġadin* is attested in Kor 59, 18. The contexts are linguistically similar in that both spots are followed with the phrase *wa-llāhu ḫabīrun bi-mā taʿmalūn*, an expression that occurs in the Qurʾān in only five other spots. Note also the relative probable proximity of this attestation to the spot at hand. (Incidentally, the word *ġadan* occurs in the Qurʾān also in Kor 12, 12; 18, 23; 31, 34; 54, 26; and 59, 18.)

62, 6	أنكم أولياء لله	أنكم أولياء لله في الأخرة	addition (*fī l-āḫirati*)

C-1 has *fī l-āḫirati* ("in the next world"), but ʿUṯmān does not. The present verse and the next are close parallels of the pair of verses 94 and 95 in *sūra* 2. Much of the wording is almost exactly the same. One key difference is that Kor 2, 94 mentions the next world (*l-dāru l-āḫiratu*), but the present verse does not. Therefore, C-1's *fī l-āḫirati* possibly harks back to Kor 2, 94-5. Cf. the parallel 3, 142-3, which also concerns the next world but does not employ the term *āḫira*.

The phrase *fī l-āḫirati* occurs twenty-four times in the Qurʾān, and *āḫira* occurs 113 times.

(c) *Substitutions without phonetic conservation.*

Verse	ʿUṯmān	C-1	Type of Variant
2, 196	فمن كان منكم مريضا	فان كان أحد منكم مريضا	substitution (syntactical structure: *in… aḥad* vs. *man*) and phonetic conservation (same three words used in different syntactical structures).

(*cont.*)

Verse	ʿUṯmān	C-1	Type of Variant

The syntactical structures used by C-1 and ʿUṯmān occur elsewhere in the Qurʾān.

On the one hand, the language used by C-1 is quite familiar. The C-1 phrase *aḥadun minkum* and the construct *in kāna(t)* are both attested in legal passages in ʿUṯmān: The C-1 construct *in kāna(t)* occurs about forty times in ʿUṯmān, including a dozen times as part of a legal prescription: Kor 2, 280; 2, 282; 4, 11 (three times), 4, 12 (three times), 4, 92 (twice), and 4, 102. The phrase *aḥadun min* occurs three times in ʿUṯmān, twice in the form of *aḥadun minkum*. The occurrences are all by way of legal prescription (Kor 43, 29; 5, 6; 9, 6).

On the other hand, ʿUṯmān's version of this verse, with *man kāna minkum marīḍan*, too, has several close parallels as part of a legal prescription: Kor 2, 185; 2, 196; 2, 232; 4, 6 (twice). The sequence *man kāna* occurs thirty-seven times in ʿUṯmān, usually in a non-legal passage.

For another instance of C-1's propensity to construct conditionals with *in* instead of *man*, see below, Kor 5, 45.

Verse	ʿUṯmān	C-1	Type of Variant
2, 197	تفعلوا	تعملوا	substitution (*tafʿalū* vs. *taʿmalū*)

The two words are graphically similar, so the difference might be due to copying from manuscript.

ʿUṯmān's *mā tafʿalū min ḫayrin* is attested also in Kor 2, 215 and 4, 127.
The synonyms are both very common verbs in the Qurʾān. The verb *faʿala* of ʿUṯmān occurs in the Qurʾān 108 times, while the verb *ʿamala* of C-1 occurs some 359 times in different conjugations.

Verse	ʿUṯmān	C-1	Type of Variant
2, 209	جاتكم البينت	جاءكم الهدى	substitution (*ǧāʾatkumu l-bayyināt* vs. *ǧāʾakumu l-hudā*)

Both phrases occur elsewhere. We have *ǧāʾahumu l-hudā* in Kor 17, 94 and 18, 55 (cf. 28, 37 and 28, 85). There is likewise *ǧāʾakum bayyinatun* in Kor 6, 157; 7, 73; and 7, 85 (cf. 40, 28).

(*cont.*)

Verse	ʿUṯmān	C-1	Type of Variant
2, 213	فبعث	فأرسل	substitution: synonyms (*fa-baʿaṯa* vs. *fa-arsala*)

The verb *baʿaṯa* occurs sixty-five times in the Qurʾān, mostly in connection with resurrection. The verb *arsala* occurs 172 times. While both words can be used for the sending of prophets (*i.e.* the subject of the verse at hand), C-1's *arsala* is used significantly more frequently in such a context.

5, 46	نور ومصدقا لما بين يديه من التورية	نور ومصدقا لما أنزلنا من التورية	substitution (*bayna yadayhi* vs. *anzalnā*)

ʿUṯmān's expression *bayna yaday* (or *bayna aydī*) occurs forty times in the Qurʾān. Close parallels would include Kor 2, 97; 3, 3; 3, 50; 5, 46 (another instance), 5, 48; 6, 92; 10, 37; 12, 111; 35, 31; 36, 9; 46, 30, and 61, 6.

As for the C-1 expression, the verb *anzala* is juxtaposed with *al-tawrāt* in Kor 3, 3; 3, 65, and 5, 44, in the form of *anzalnā* in the first and last cases. The verb *anzalnā* occurs fifty-five times in the Qurʾān.

In sum, the expressions are both attested elsewhere, and are both attested in close proximity to the point at hand.

5, 46	للمتقين	لقوم (يومنون or يوقنون)	substitution (*li-l-muttaqīna* vs. *li-qawmin yuʾminūna* or *li-qawmin yūqinūna*).

ʿUṯmān has *mawʿiẓatan li-l-muttaqīna*, a phrase that occurs in three other spots, *viz.* Kor 2, 66; 3, 138; and 24, 34. (The word *li-l-muttaqīna* occurs seventeen other times in the Qurʾān.)

The wording in C-1 is *mawʿiẓatan li-qawmin yuʾminūna*, or, possibly, *li-qawmin yūqinūna*. Elsewhere, the word *mawʿiẓatan* is never followed by such a phrase. The phrase *li-qawmin yuʾminūna* occurs in thirteen spots: Kor 6, 99; 7, 52; 7, 188; 7, 203; 12, 111; 16, 64; 16, 79; 27, 86; 28, 3; 29, 24; 29, 51; 30, 37; 39, 52. The word *yūqinūna* occurs eleven times, including four times within the phrase *li-qawmin yūqinūna*: Kor 2, 118; 5, 50; 45, 4; 45, 20. Note that the instance at Kor 5, 50 is close to the spot at hand and one could imagine it having influenced the phrasing of the verse at hand.

(*cont.*)

Verse	ʿUṯmān	C-1	Type of Variant
5, 49	أنزل الله	أوحى الله	substitution (synonyms *anzala* vs. *awḥā*)

The verb in ʿUṯmān, *anzala,* occurs 190 times in different senses and conjugations, including twenty-nine times as *anzala llāhu.* The term is used not only with revelations, but also in other contexts.

The verb in C-1, *awḥā,* occurs seventy-three times in various conjugations, including eight times as *awḥā.* The term is used specifically with revelation.

Neither term could be considered unexpected in context.

Verse	ʿUṯmān	C-1	Type of Variant
63, 4	وإذا رأيتهم	وإذا (جاؤك)	substitution (*raʾaytahum* instead of *ǧāʾūka*)

C-1's جاءوك harks back to جاءك in the first verse of the *sūra.* The word *ǧāʾaka* occurs eleven times in the Qurʾān, including three times following *iḏā.* See above, Kor 5, 42.

ʿUṯmān's *raʾaytahum* occurs elsewhere three times (Kor 20, 92; 33, 19; 76, 19), but never in a similar context. The closest parallel would be Kor 76, 19.

(d) *Errors of the Hand.*

Verse	ʿUṯmān	C-1	Type of Variant
2, 196	ففدية	فدية	minor omission (*fa*) probable error of the hand
2, 200	فإذا قضيتم	وإذا أقضيتم	scribal error; duplication of *alif*
2, 202	كسبوا	كسبون	scribal error: *nūn* instead of *alif* For the same type of error, see below, Kor 2, 217.

(*cont.*)

Verse	ʿUṯmān	C-1	Type of Variant
2, 213	ليحكم بين الناس	ليحكم (ا)بين الناس	scribal error: addition of *alif*
2, 217	إن استطعوا	إن استطعو (ن)	scribal error: *nūn* instead of *alif* For the same type of error, see above, Kor 2, 202.
63, 7	ولكن	ولكن (ا)ن	scribal error

(e) *Illegible terms that differ from ʿUṯmān.*

Verse	ʿUṯmān	C-1	Type of Variant
5, 54	أذلة	؟؟؟	Illegible
5, 54	أعزة	؟؟؟	Illegible
15, 68	إن هؤلاء	؟	Illegible
89, 27	النفس المطمئنة	النفس (–هـ–) المطمئنة	illegible; error of the hand?
89, 28	ارجعي إلى	ا – ؟	illegible (there is room for only one or two more letters after the *alif*)

Appendix 3: *Sūra* Titles in C-1.

Items placed in parentheses (x) are only partly visible, but enough is visible to have good reason for the readings given here.

Verse	C-1	Note
63 End	هذه خىمة سـ(ـورة) المنفقين	*sūra* title
62 End	(هذه) خىمة سورة الجمـ(ـعة)	*sūra* title
89 End	هـ (ذه) خىمة اا (فجر) وليل العشر	*sūra* title

Appendix 4: Orthographical Characteristics of C-1

Items placed in parentheses (x) are only partly visible, but enough is visible to have good reason for the readings given here. Although dots and *hamzas* are used below, they tend not to appear in the original.

Verse	(Modern Saʿūdī Qurʾān)	C-1
2, 193	حتى	حتا
2, 194	الحرام...الحرام	الحرم...الحرم
2, 194	قصاص	قصص
2, 196	حتى	حتا
2, 196	من صيام	من صيم
2, 196	فصيام	فصيم
2, 196	كاملة	كملة
2, 196	حاضرى	حضرى
2, 196	الحرام	الحرم
2, 196	جدال	جدل
2, 197	الزاد	الزد
2, 197	يأولى الألبب	يأولا الألبب
2, 198	جناح	جنح
2, 198	الحرام	الحرم
2, 198	الضالين	الضلين
2, 199	أفاض	أفض
2, 203	أيام	أيم
2, 204	الخصام	الخصم
2, 206	المهاد	المهد
2, 207	مرضات	(مرضت)
2, 207	بالعباد	بالعبد
2, 210	الغمام	الغمم
2, 210	أتينهم	أتينيهم
		(the long *ā* designated with a tooth, no dots)

(*cont.*)

Verse	(Modern Saʿūdī Qurʾān)	C-1
2, 211	العقاب	العقب
2, 212	حساب	(حسب)
2, 214	الباسا	البسا
2, 216	القتال	القتل
2, 217	الحرام	الحرم
2, 217	قتال	قتل
2, 217	قتال	قتل
2, 217 (verso)	لايزالون	لايزلون
2, 217	يقاتلونكم	يقتلونكم
2, 218	هاجروا	هجروا
2, 218	رحمت	رحمة
2, 219	يسئلونك	يسألونك
2, 219	ويسئلونك	ويسألونك
2, 220	ويسئلونك	ويسألونك
2, 220	إصلاح	إصلح
2, 220	تخالطوهم	تخلطوهم
2, 222	ويسئلونك	ويسألونك
5, 44	هادوا	هدوا
5, 44	بأيتي	بأيئتي without diacritics (just extra tooth)
5, 48	عما	عن ما
5, 48	منهاجا	منهجا
5, 48	في ما	فيما
5, 52	دائرة	دئرة
15, 55	الضالون	الضلون
15, 58	قالوا	قلوا
15, 71	قال	قل
63, 4	أجسامهم	أجسمهم

(*cont.*)

Verse	(Modern Saʿūdī Qurʾān)	C-1
63, 5	تعالوا	(تعلوا)
63, 5	رأيتهم	ريتهم
63, 7	حتي	حتا
63, 7	خزائن	خزئن
62, 5	الحمار	الحمر
62, 5	أسفارا	أسفرا
62, 6	هادوا	هدوا
89, 7	العماد	العمد
89, 8	البلد	(البليد) just a tooth, no dots, for *ā*
89, 9	جابوا	جيوا
89, 10	الأوتاد	الأوتد
89, 11	البلد	(البليد) just a tooth, no dots, for *ā*
89, 14	لبالمرصاد	لبالمرصد
89, 18	طعام	طعم
89, 26	وثاقه	وثقه
89, 28	راضية	رضية